

POLSKA AKADEMIA NAUK  
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

PLISSN 00- 0035-938

KWARTALNIK  
ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND  
TELECOMMUNICATIONS  
QUARTERLY

dawniej  
ROZPRAWY ELEKTROTECHNICZNE

TOM XXXVI - ZESZYT 1

PAŃSTWOWE WYDAWNICTWO NAUKOWE  
WARSZAWA 1990



POLSKA AKADEMIA NAUK  
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

KWARTALNIK  
ELEKTRONIKI I TELEKOMUNIKACJI  
(ELECTRONICS AND  
TELECOMMUNICATIONS  
QUARTERLY)

TOM XXXVI — ZESZYT 1

PAŃSTWOWE WYDAWNICTWO NAUKOWE  
WARSZAWA 1990

## RADA REDAKCYJNA

### *Przewodniczący*

prof. dr ADAM SMOLIŃSKI

członek rzeczywisty PAN

### *Członkowie*

prof. dr hab. DANIEL JÓZEF BEM, prof. dr hab. inż. MICHAŁ BIAŁKO — czł. koresp. PAN,  
prof. dr inż. **KAZIMIERZ DZIĘCIOŁOWSKI**, prof. dr hab. inż. STEFAN HAHN — czł. koresp.  
PAN, prof. dr inż. ANDRZEJ HAŁAS, prof. dr inż. ZDZISŁAW KACHLICKI, prof. dr hab. inż.  
BOHDAN MROZIEWICZ, prof. dr inż. JERZY OSIOWSKI, prof. dr inż. WITOLD ROSIŃSKI — czł.  
rzecz. PAN, prof. dr inż. STANISŁAW SŁAWIŃSKI, prof. dr hab. inż. STEFAN WĘGRZYN — czł.  
rzecz. PAN, prof. dr hab. inż. WIESŁAW WOLIŃSKI, prof. dr inż. ANDRZEJ ZIELIŃSKI,  
prof. dr inż. MARIAN ZIENTALSKI

## REDAKCJA

### *Redaktor naczelny*

prof. dr hab. inż. WIESŁAW WOLIŃSKI

### *Zastępca Redaktora Naczelnego*

doc. dr inż. KRYSTYN PLEWKO

### *Sekretarz Odpowiedzialny*

mgr KRYSTYNA LELAKOWSKA

## ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika,  
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

*Dyżury Redakcji: środy i piątki, godz. 14—16*

*tel. 28 89 81 210 07-737*

*Telefony domowe: Redaktora Naczelnego: 12 17 65*

*Zast. Red. Naczelnego: 26 83 41*

*Sekretarza Odpowiedzialnego: 25 29 18*

Państwowe Wydawnictwo Naukowe

Warszawa, ul. Miodowa 10

|                                         |                                         |
|-----------------------------------------|-----------------------------------------|
| Ark. wyd. 16,5                          | Oddano do składania we wrześniu 1990 r. |
| Ark. druk. 14,5                         | Podpisano do druku w kwietniu 1991 r.   |
| Papier offsetowy kl. III 70 g, 70 × 100 | Zamówienie nr 1636/12/90                |

Drukarnia Naukowo-Techniczna, Warszawa

## Szanowni Czytelnicy

Zeszyt naszego czasopisma przekazywany w Wasze ręce jest pierwszym, który ukazuje się pod nowym tytułem

# KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI (ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY)

Kwartalnik ten jest bezpośrednią kontynuacją wydawanego od 36 lat pod auspicjami Polskiej Akademii Nauk kwartalnika pt. „Rozprawy Elektrotechniczne”.

„Kwartalnik Elektroniki i Telekomunikacji” jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk i w związku z tym na jego łamach publikowane są prace naukowe dotyczące podstaw teoretycznych i zastosowań z zakresu elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej. Zamieszczane są prace oryginalne, w tym przyczynkowe oraz zawierające analizy porównawcze metod lub systemów, systematyczne ujęcia określonej dziedziny wiedzy, omówienia aktualnego stanu lub postępu danej gałęzi techniki oraz omówienia perspektyw rozwojowych.

Czasopismo jest przeznaczone dla specjalistów oraz studentów zajmujących się wymienioną tematyką. Artykuły są publikowane w języku polskim i angielskim.

Redakcja wyraża ubolewanie, że z przyczyn od nas niezależnych poszczególne zeszyty kwartalnika ukazują się z rocznym opóźnieniem. Bardzo serdecznie za to przepraszamy naszych Autorów i Czytelników. Mamy nadzieję, że ten stan ulegnie poprawie.

Opóźnienia w wydawaniu kwartalnika powodują zakłócenia w jego zamawianiu i dostarczaniu do naszych Czytelników. W związku z tym informujemy, że:

- numery kwartalnika „Rozprawy Elektrotechniczne” z poprzednich lat i numery „Kwartalnika Elektroniki i Telekomunikacji” od 1990 r. można nabyć w Księgarni Państwowego Wydawnictwa Naukowego w Warszawie, ul. Miodowa 10.
- „Kwartalnik Elektroniki i Telekomunikacji” można prenumerować i kupować w Ośrodku Rozpowszechniania Wydawnictw Naukowych Polskiej Akademii Nauk

### ORPAN

w Warszawie: Pałac Kultury i Nauki  
00-901 Warszawa, centrala tel. 20-02-11 wew. 21-05 Dział Wydawnictw Krajowych

w Białymstoku: 15-082 Białystok, ul. Nowotki 13 tel. 415-070 wew. 175

w Lublinie: 20-031 Lublin, Pl. Marii Skłodowskiej-Curie 5 tel. 375-413

w Łodzi: 90-265 Łódź, ul. Piotrkowska 48 tel. 333-702

w Katowicach: 40-007 Katowice, ul. Bankowa 11 tel. 587-709

w Krakowie: 31-020 Kraków, ul. Św. Marka 22 tel. 215-314

w Poznaniu: 61-725 Poznań, ul. Mielżyńskiego 27/29 tel. 524-516

we Wrocławiu: 50-071 Wrocław, Pl. Wolności 7 tel. 340-16

Cena poszczególnych zeszytów „Kwartalnika Elektroniki i Telekomunikacji” wynosi 10 000 zł za numer.

W przypadkach trudności w zaprenumerowaniu lub nabyciu zeszytów kwartalnika prosimy o informowanie o tym naszej redakcji.

REDAKCJA



# Nieliniowe elementy dynamiczne w aspekcie realizacji praktycznej

STANISŁAW OSOWSKI

*Instytut Elektrotechniki Teoretycznej i Miernictwa Elektrycznego.  
Politechnika Warszawska*

*Otrzymano 1988.12.7*

*Autoryzowano do druku 1989.4.3*

W pracy przedstawiono metody realizacji elementów i złożonych układów dynamicznych z wykorzystaniem teorii obwodów rezystancyjnych. Poddano badaniom wybrane układy realizujące mutatory różnych typów w aspekcie ich zastosowania do symulacji elementów reaktancyjnych. Dokonano przeglądu metod transformacji opisu z przestrzeni stanu do postaci obwodowej odpowiadającej równaniom stanu.

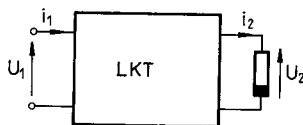
## 1. WPROWADZENIE

W teorii obwodów nieliniowych wyróżnia się cztery podstawowe rodzaje elementów nieliniowych związane z istnieniem czterech zmiennych: prądu  $i$ , napięcia  $u$ , strumienia  $\Psi$  i ładunku  $q$ . Są to: rezystor opisany w pełni przy użyciu wektora prądu  $i$  oraz napięcia  $u$  za pomocą równania algebraicznego  $f(u, i, t) = 0$ , kondensator scharakteryzowany równaniem  $f(u, q, t) = 0$  oraz  $i = dq/dt$ , indukcyjność opisana za pomocą równania  $f(i, \psi, t) = 0$  przy  $u = d\psi/dt$  a także rezystor wiążący w swoim opisie strumień i ładunek  $f(q, \psi, t) = 0$  przy czym  $i = dq/dt$  oraz  $u = d\psi/dt$ . Trzy ostatnie elementy nieliniowe kondensator, indukcyjność oraz rezystor tworzą rodzinę elementów dynamicznych (reaktancyjnych) albowiem do ich pełnego opisu oprócz podstawowej zależności algebraicznej typu  $f(x, y) = 0$  wymagana jest jeszcze zależność różniczkowa wiążąca prąd i napięcie odpowiednio z ładunkiem i strumieniem. Jak zostało wykazane w pracach [1, 2, 3] teoria obwodów rezystancyjnych jest podstawą realizacji układów dynamicznych gdyż każdy element dynamiczny może być zrealizowany poprzez odpowiednie połączenie elementów nieliniowych rezystancyjnych oraz elementów reaktancyjnych i źródeł sterowanych.

W pracy przedstawiony zostanie przegląd obwodów transformacyjnych pozwalających zrealizować dowolny element dynamiczny. Dokonana będzie analiza tych rozwiązań włącznie z ich weryfikacją komputerową. Rozważony zostanie przypadek bardziej rozbudowanych obwodów dynamicznych zawierających różne rodzaje elementów nieliniowych w aspekcie ich realizacji praktycznej. Przedstawione zostaną przykłady liczbowe ilustrujące rozważania ogólne.

## 2. REALIZACJA ELEMENTÓW DYNAMICZNYCH

Element dynamiczny dowolnego typu można zrealizować przy pomocy rezystora nieliniowego i liniowego konwertera transformacyjnego (LKT) w obwodzie jak na rys. 1 [1, 3, 4, 5]. Zasada przekształcania elementu rezystancyjnego na typ dynamiczny polega na tym, że wielkości prądu i napięcia  $i_2$ ,  $u_2$  na zaciskach wtórnych czwórnika



Rys. 1. Ogólna struktura obwodu realizującego element dynamiczny z wykorzystaniem LKT

ulegają przekształceniu na inne wielkości, włącznie ze zmianą typu (strumień  $\psi_1$ , ładunek  $q_1$  bądź prąd  $i_1$  i napięcie  $u_1$ ) na zaciskach pierwotnych. Cel ten może zrealizować tylko czwórnik o parametrach uzależnionych częstotliwościowo. Najczęściej do tego celu używa się bądź to konwertera impedancyjnego o macierzy łańcuchowej  $A$

$$A = \begin{bmatrix} K_1(s) & 0 \\ 0 & K_2(s) \end{bmatrix}, \quad (1)$$

bądź też inwertera impedancyjnego o opisie łańcuchowym  $A$

$$A = \begin{bmatrix} 0 & K_1(s) \\ K_2(s) & 0 \end{bmatrix}. \quad (2)$$

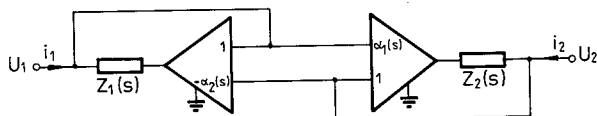
W przypadku inwertera impedancyjnego istnieje również opis admitancyjny wygodniejszy w użyciu

$$Y = \begin{bmatrix} 0 & K_2(s) \\ -1 & 0 \\ K_1(s) & \end{bmatrix}. \quad (3)$$

Oba typy czwórników noszą ogólną nazwę mutatorów [1, 3, 4]. Ze względu na stabilność realizowanego czwórnika jak również łatwość realizacji przyjmuje się najczęściej  $K_1(s)$  i  $K_2(s)$  jako funkcje typu  $ks$  bądź  $k/s$  lub jako wartość stałą. Oba typy czwórników transformacyjnych są czwórnikami liniowymi a szczególne ich rodzaje jak np. GIC są szeroko znane i stosowane w teorii układów liniowych. Jednakże ich zastosowanie w realizacji elementów dynamicznych nieliniowych na bazie elementów rezystancyjnych jest zwykle zbyt ograniczone aby mogło być brane pod uwagę jako rozwiązanie problemu. Jakkolwiek istnieją opracowane jednolite struktury ogólne realizujące dowolny typ mutatora [4] wydaje się, że z praktycznego punktu widzenia (wrażliwość, stabilność) należy oddzielnie podejść do realizacji konwertera i inwertera impedancyjnego.

## 2.1. INWERTER IMPEDANCYJNY

W przypadku inwertera impedancyjnego odpowiednią strukturą obwodu jest realizacja przedstawiona na rys. 2 będąca szczególnym przypadkiem ogólnej struktury czwórnika opracowanej w [4, 5].



Rys. 2. Struktura obwodowa inwertera impedancyjnego

Jak łatwo udowodnić, przy założeniu idealności elementów, macierz admitancyjna tego obwodu wyraża się następująco

$$Y = \begin{bmatrix} 0 & Y_1 \alpha_2(s) \\ -Y_2 \alpha_1(s) & 0 \end{bmatrix}. \quad (4)$$

Z porównania macierzy (3) i (4) wynika, że

$$K_1(s) = \frac{1}{Y_2 \alpha_1(s)} \quad (5)$$

$$K_2(s) = Y_1 \alpha_2(s). \quad (6)$$

Stąd realizacja funkcji  $K_1(s)$  i  $K_2(s)$  następuje przez selekcję admitancji  $Y_1$  i  $Y_2$  oraz dobór odpowiednich wzmocnień  $\alpha_1(s)$  i  $\alpha_2(s)$  sumatora. Dla przykładu realizując  $K_1(s) = \gamma_1 s$  oraz  $K_2(s) = \gamma_2 s$  można przyjąć  $\alpha_1(s) = k_1/s$ ,  $\alpha_2(s) = k_2$ ,  $Y_1 = s\gamma_2/k_2$  oraz  $Y_2 = 1/(k_1\gamma_1)$  gdzie  $k_1$  i  $k_2$  są współczynnikami skali uzależnionymi od konkretnej realizacji. Tablica 1 przedstawia przykładowe rozwiązania mutatorów przekształcających nieliniową rezystancję obciążenia w odpowiedni typ elementu dynamicznego (indukcyjność, pojemność, rezystancję) przy czym  $k_1$  i  $k_2$  są dobieranymi przez użytkownika współczynnikami skali. W przedstawionym w tablicy rozwiązaniu przyjęto ze względów praktycznych, że wzmacniacz sumacyjny sumuje sygnały ze

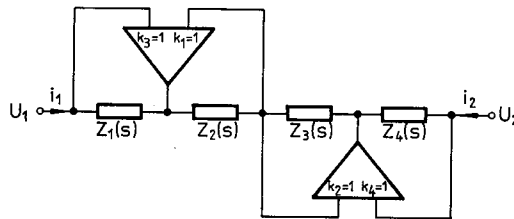
Tablica 1

| Typ przekształcenia | Macierz A                                                        | Transformacja zmiennych                           | Parametry obwodu |               |                          |                          |
|---------------------|------------------------------------------------------------------|---------------------------------------------------|------------------|---------------|--------------------------|--------------------------|
|                     |                                                                  |                                                   | $\alpha_1(s)$    | $\alpha_2(s)$ | $Y_1$                    | $Y_2$                    |
| $\psi \Leftarrow R$ | $\begin{bmatrix} 0 & \gamma_1 s \\ \gamma_2 & 0 \end{bmatrix}$   | $i_1 \Rightarrow \psi_1$<br>$u_2 \Rightarrow i_1$ | $\frac{k_1}{s}$  | $k_2$         | $\frac{\gamma_2}{k_2}$   | $\frac{1}{k_1 \gamma_1}$ |
| $q \Leftarrow R$    | $\begin{bmatrix} 0 & \gamma_1 \\ \gamma_2 s & 0 \end{bmatrix}$   | $i_2 \Rightarrow u_1$<br>$u_2 \Rightarrow q_1$    | $k_1$            | $k_2$         | $\frac{\gamma_2 s}{k_2}$ | $\frac{1}{\gamma_1 k_1}$ |
| $M \Leftarrow R$    | $\begin{bmatrix} 0 & \gamma_1 s \\ \gamma_2 s & 0 \end{bmatrix}$ | $i_2 \Rightarrow \psi_1$<br>$u_2 \Rightarrow q_1$ | $\frac{k_2}{s}$  | $k_2$         | $\frac{\gamma_2 s}{k_2}$ | $\frac{1}{\gamma_1 k_1}$ |

wzmocnieniem bądź rzeczywistym bądź całkując je. Inne rozwiązanie wykorzystujące operacje różniczkowania jest też możliwe ale mało praktyczne ze względu na dużą wrażliwość układu na szумы. Należy zwrócić uwagę że mutator typu inwertera dokonuje transformacji zmiennych z przekształceniem skali prądowej w napięciową i odwrotnie, np. w przypadku indukcyjności nieliniowej napięcie  $u_2$  przekształcane jest w prąd wejściowy  $i_1$  natomiast prąd  $i_2$  w strumień  $\psi_1 = \int u_1 dt$ . Stąd transformacja taka pokrywa jedynie jedną klasę zagadnień np. rezystor nieliniowy uzależniony napięciowo będzie transformowany w element dynamiczny uzależniony bądź prądowo bądź ładunkowo ale nigdy uzależniony napięciowo bądź strumieniowo. Dla uzyskania pełnej transformacji należy wprowadzić drugą komplementarną klasę mutatorów typu konwertera.

## 2.2. KONWERTER IMPEDANCYJNY

Ogólna struktura obwodu realizującego dowolny typ konwertera impedancyjnego opisanego równaniem (1) podana jest na rys. 3. Obwód zawiera cztery impedancje  $Z_i$  ( $i = 1, 2, 3, 4$ ) oraz dwa wzmacniacze sumacyjne o wzmocnieniach jednostkowych.



Rys. 3. Struktura obwodowa konwertera impedancyjnego

Przyjmując wzmacniacze jako idealne, tzn. wzmocnienia  $+1$ , nieskończona impedancja wejściowa, zerowa impedancja wyjściowa łatwo udowadnia się, że układ przedstawiony na rys. 3 opisuje się macierzą łańcuchową o postaci

$$A = \begin{bmatrix} -Z_2 Y_3 & 0 \\ 0 & -Y_1 Z_4 \end{bmatrix}, \quad (7)$$

przy czym  $Y_1 = 1/Z_1$ ,  $Y_3 = 1/Z_3$ . Jak z powyższego wynika opis (7) odpowiada macierzy konwertera impedancyjnego (1) przy czym  $K_1(s) = -Z_2 Y_3$ ,  $K_2(s) = -Y_1 Z_4$ . Transformacja prądu i napięcia dokonuje się ze zmianą znaku co w praktyce oznacza odwrócenie fazy odpowiednich wielkości. Zmiana znaku zarówno prądu jak i napięcia nie wpływa na immitancję wejściową widzianą od strony wejścia. Realizacja funkcji  $K_1(s)$  oraz  $K_2(s)$  odbywa się poprzez odpowiedni dobór impedancji  $Z_i$  ( $i = 1, 2, 3, 4$ ) elementów pasywnych.

W tablicy 2 zebrano przykładowe rozwiązania konwertera przekształcającego nieliniową rezystancję obciążenia w odpowiedni typ elementu dynamicznego (indukcyjność, pojemność, rezystancja).

Zauważmy, że w poszczególnych przypadkach przedstawionych w tablicy 2 relacje między parametrami macierzy  $A$  i obwodu są następujące: dla indukcyjności

Tablica 2

| Typ przekształcenia | Macierz $A$                                                      | Transformacja zmiennych                           | Parametry obwodu |       |        |       |
|---------------------|------------------------------------------------------------------|---------------------------------------------------|------------------|-------|--------|-------|
|                     |                                                                  |                                                   | $Y_1$            | $Z_2$ | $Y_3$  | $Z_4$ |
| $\psi \Leftarrow R$ | $\begin{bmatrix} \gamma_1 s & 0 \\ 0 & \gamma_2 \end{bmatrix}$   | $u_2 \Rightarrow \psi_1$<br>$i_2 \Rightarrow i_1$ | $G_1$            | $R_2$ | $sC_3$ | $R_4$ |
| $q \Leftarrow R$    | $\begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 s \end{bmatrix}$   | $u_2 \Rightarrow u_1$<br>$u_2 \Rightarrow q_1$    | $sC_1$           | $R_2$ | $G_3$  | $R_4$ |
| $M \Leftarrow R$    | $\begin{bmatrix} \gamma_1 s & 0 \\ 0 & \gamma_2 s \end{bmatrix}$ | $u_2 \Rightarrow \psi_1$<br>$i_1 \Rightarrow q_1$ | $sC_1$           | $R_2$ | $sC_3$ | $R_4$ |

$\gamma_1 = R_2 C_3$ ,  $\gamma_2 = G_1 R_4$ , dla pojemności  $\gamma_1 = R_2 G_3$ ,  $\gamma_2 = C_1 R_4$  natomiast dla rezystancji  $\gamma_1 = R_2 C_3$  i  $\gamma_2 = C_1 R_4$ .

Istotną zaletą układu jest prostota w zamianie jednego typu konwertera na drugi poprzez wymianę jedynie niektórych ( $Y_1$ ,  $Y_3$ ) elementów pasywnych. Dostrajanie obwodu w celu realizacji odpowiednich funkcji  $K_1(s)$  i  $K_2(s)$  jest niezależne od siebie i odbywa się poprzez dobór odpowiedniej wartości elementów pasywnych, przy czym elementy te nie wpływają na siebie wzajemnie.

Pewną wadą rozwiązania jest konieczność starannego doboru wartości wzmocnień wzmacniaczy sumacyjnych. W pełni diagonalna macierz  $A$  powstaje jedynie w przypadku wszystkich wzmocnień równych jedności. Jeśli warunek ten nie jest zachowany, w macierzy  $A$  pojawiają się elementy resztkowe poza diagonalną które zmniejszają dobroć realizowanego elementu, a w szczególnych przypadkach mogą stać się powodem wzbudzenia drgań pasożytniczych. Dla przykładu jeśli zamiast wzmocnienia jednostkowego  $k_1$  i  $k_2$  zostaną zrealizowane inne wartości, wówczas macierz  $A$  wyrazi się następująco

$$A = \begin{bmatrix} -Z_2 Y_3 & -\frac{Z_4}{k_2} (1 - k_1 + Z_2 Y_3 - k_2 Z_2 Y_3) \\ 0 & \frac{-k_1}{k_2} Y_1 Z_4 \end{bmatrix} \quad (8)$$

Jedynie dla  $k_1 = k_2 = 1$  macierz ta przyjmuje postać pożądaną (7). W przeciwnym wypadku element  $A_{12}$  macierzy  $A$  jest różny od zera. Realizując np. liniową indukcyjność  $L$  w układzie konwertera obciążonego rezystorem  $R_0$  przy doborze jego

elementów zgodnie z pierwszą pozycją tablicy 2 otrzymuje się wyrażenie na impedancję wejściową w postaci

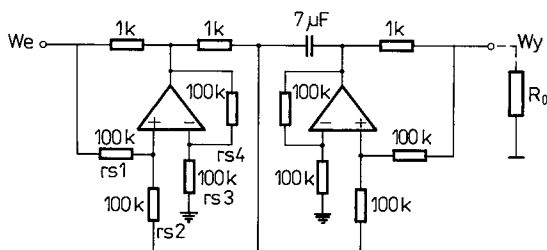
$$Z_{we} = s \left[ \frac{k_2 R_2 C_3}{k_1 G_1 R_4} R_0 + \frac{k_2 R_2 C_3}{k_1 G_1} - \frac{R_2 C_3}{G_1} \right] + \frac{k_1 - 1}{k_1 G_1}. \quad (9)$$

Jedynie dla  $k_1 = k_2 = 1$  zrealizowana zostanie idealna indukcyjność  $L = k_2 R_2 C_3 R_0 / (k_1 G_1 R_4)$ , w przeciwnym wypadku indukcyjność jest stratna, a rezystancja tych strat  $R_1 = (k_1 - 1) / (k_1 G_1)$  zależy od odchylenia wzmocnienia  $k_1$  od wartości jednostkowej. Z równania (9) wynika również warunek na stabilność obwodu  $k_1 \geq 1$  (przy założeniu niewielkiej odchyłki  $k_1$  i  $k_2$  od wartości znamionowej). Jeśli zrealizowana wartość  $k_1$  jest mniejsza od jedności rezystancja  $R_1$  staje się ujemna a obwód niestabilny.

### 3. SYMULACJA KOMPUTEROWA UKŁADÓW

Dla potwierdzenia słuszności rozważań teoretycznych została przeprowadzona symulacja wybranych rozwiązań mutatorów realizujących elementy dynamiczne. Poddano tym badaniom zarówno układ inwertera jak i konwertera impedancji uzyskując zadowalające rezultaty. Dla potwierdzenia użyteczności praktycznej rozwiązania dokonano symulacji obwodów z rzeczywistymi elementami aktywnymi. Jakość tych wyników i stopień zbieżności z oczekiwaniami teoretycznymi w znacznej mierze zależał od przyjętych parametrów wzmacniacza operacyjnego. Jedynie dla nieskończonego wzmocnienia wzmacniaczy wyniki zarówno dla analizy czasowej jak i częstotliwościowej idealnie pokrywały się z oczekiwaniami. Przy uwzględnieniu parametrów rzeczywistych zgodność ta występowała w ograniczonym zakresie częstotliwości do 10 kHz.

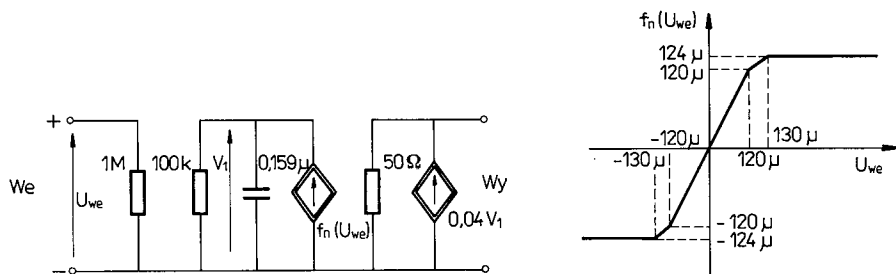
Rys. 4 przedstawia stosowany w badaniach układ konwertera impedancyjnego przekształcającego rezystancję obciążenia w indukcyjność. Jak z rysunku widać wzmacniacze sumacyjne budowane były na bazie wzmacniacza operacyjnego ze



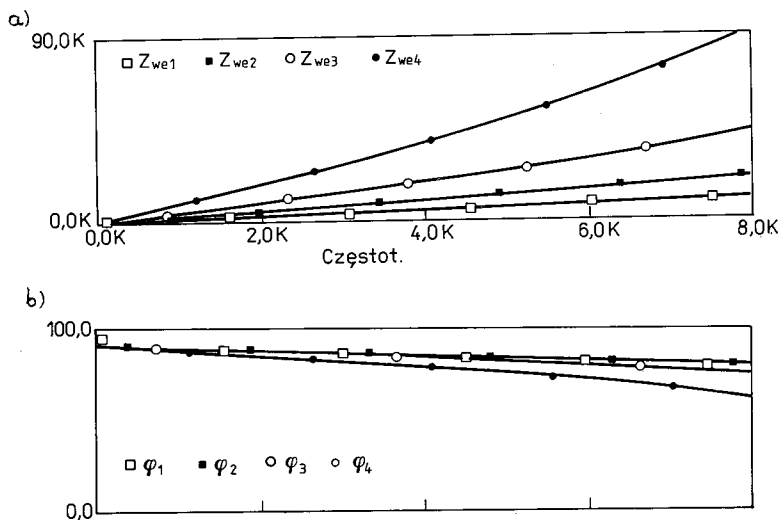
Rys. 4. Schemat konwertera poddanego badaniom

sprężeniem rezystancyjnym o rezystancji  $R = 100 \text{ k}\Omega$ . Układ był obciążany po stronie wyjściowej zarówno rezystancją liniową jak i nieliniową. W symulacji komputerowej przyjęto makromodel wzmacniacza operacyjnego o charakterystyce częstotliwościowej aproksymowanej przy pomocy jednego bieguna  $A(s) = GB/(s + \omega_0)$  o wartościach

$GB = 2 \text{ MHz}$  i  $\omega_0 = 2\pi \cdot 10 \text{ Hz}$ , rezystancji wejściowej równej  $1 \text{ M}\Omega$ , rezystancji wyjściowej  $50 \Omega$  i charakterystyce  $U_{wy} = f(U_{we})$  z nasyceniem na poziomie  $12.4 \text{ V}$  (rys. 5). Na rys. 6 przedstawiono uzyskane charakterystyki amplitudowe i fazowe impedancji wejściowej układu w funkcji częstotliwości przy obciążeniu liniowym dla różnych rezystancji obciążenia  $R_0 = 25 \Omega, 50 \Omega, 100 \Omega$  i  $200 \Omega$  w zakresie częstotliwości do  $8 \text{ kHz}$ . Zrealizowane wartości indukcyjności odpowiadające kolejnym wartościom  $R_0$  nie są wartościami stałymi w zakresie częstotliwości do  $8 \text{ kHz}$  a przebieg tych



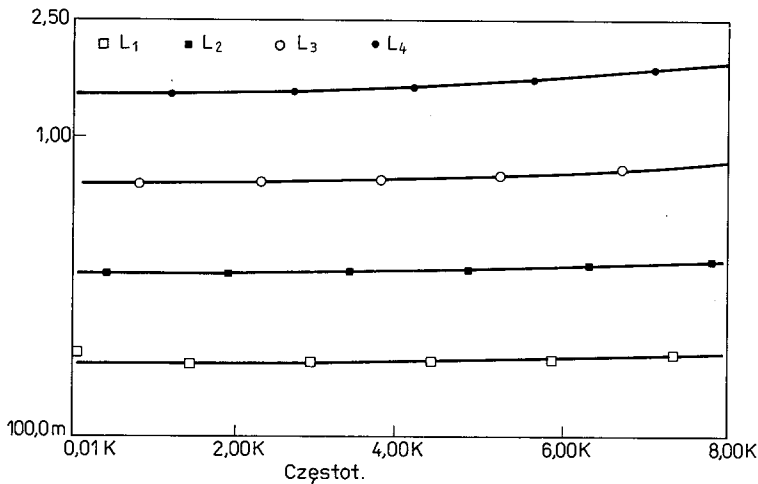
Rys. 5. Makromodel wzmacniacza operacyjnego przyjęty w badaniach



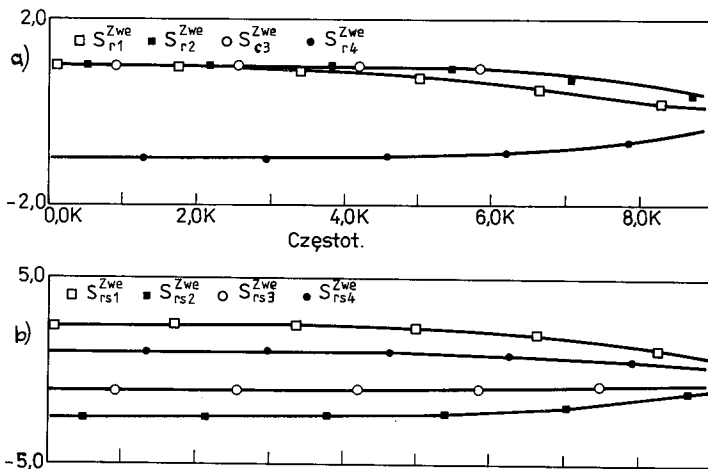
Rys. 6. Charakterystyki częstotliwościowe impedancji wejściowej konwertera obciążonego rezystorem liniowym; a) charakterystyka amplitudowa, b) charakterystyka fazowa

charakterystyk dla różnych wartości rezystancji obciążenia przedstawia rys. 7. W zakresie małych częstotliwości wartości te są praktycznie stałe i wynoszą, odpowiednio:  $0.1742 \text{ H}$ ,  $0.3488 \text{ H}$ ,  $0.6982 \text{ H}$  i  $1.398 \text{ H}$  co w sposób zadowalający odpowiada wartościom oczekiwany:  $0.175 \text{ H}$ ,  $0.350 \text{ H}$ ,  $0.700 \text{ H}$  i  $1.4 \text{ H}$ . Cechą charakterystyczną którą można było zaobserwować zarówno w symulacji komputerowej jak i badaniach laboratoryjnych była ograniczona częstotliwościowo liniowość indukcyjności zależna w dużym stopniu od parametrów obwodu i obciążenia. Przy danych konwertera jak przedstawiono na rys. 4 i 5 i wartości obciążenia równej  $200 \Omega$  poczynając od

częstotliwości ok. 5 kHz funkcja impedancji  $Z_{we}$  ulega stopniowemu odkształceniu od zależności liniowej. Im większa wartość realizowanej indukcyjności tym mniejszy jest zakres częstotliwości dla której obwód pracuje w sposób idealny. Dotyczy to również charakterystyki fazowej. Zamiast spodziewanej wartości kąta przesunięcia fazowego  $\psi = 90^\circ$  niezależnie od częstotliwości uzyskane charakterystyki fazowe dla odpowiednich rezystancji (przedstawione na rys. 6b) wykazują zmienność do około  $80^\circ$ . Badania potwierdziły, że obwód jest stosunkowo mało wrażliwy na zmiany elementów pasywnych obwodu głównego  $R_1, R_2, C_3, R_4$ . Na rys. 8a przedstawiono uzyskane zależności wrażliwości względnej impedancji wejściowej na zmianę elementów pasywnych  $R_1, R_2, C_3, R_4$  przy obciążeniu  $R_0 = 50 \Omega$  (realizacja indukcyjności  $L = 0.35 \text{ H}$ ). Są one w całym badanym zakresie częstotliwości mniejsze od jedności.



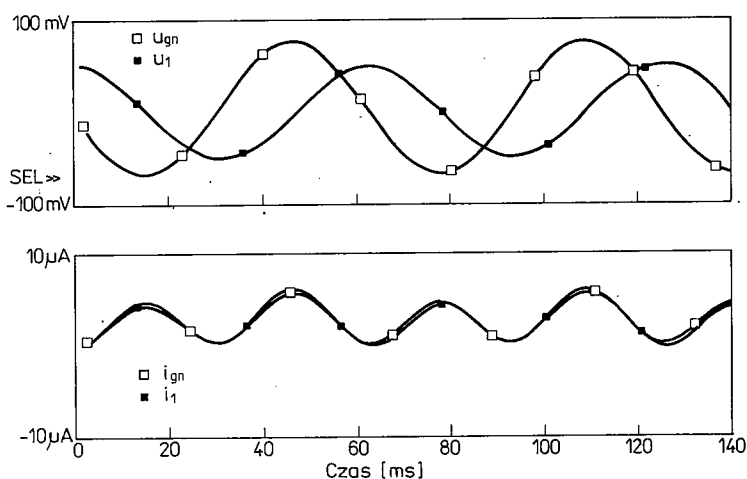
Rys. 7. Zależność zrealizowanej indukcyjności  $L$  od częstotliwości przy różnym obciążeniu rezystancyjnym obwodu



Rys. 8. Zależności wrażliwościowe badanego obwodu konwertera

W przypadku sumatorów wrażliwości obwodu są nieco większe. Na rys. 8b przedstawiono przykładowe charakterystyki wrażliwościowe względem rezystancji sumatora pierwszego. Jak z rysunku wynika są one 2—3 razy większe niż odpowiednie wielkości dla elementów pasywnych obwodu.

W badaniach nieliniowego obciążenia obwodu przyjęto obciążenie w postaci konduktancji  $G_n$  sterowanej napięciem  $G_n = 10^{-3}U_n$  i zasilanie napięciowe obwodu o wartości  $e(t) = 0.05 \cdot \cos(100 t)$ . Na rysunku 9 przedstawiono uzyskane zależności napięciowe i prądowe w obwodzie. Napięcie wejściowe obwodu zmienia się cosinusoidalnie natomiast napięcie obciążenia sinusoidalnie. W przypadku prądów następuje zgodnie z oczekiwaniami podwojenie częstotliwości przebiegu. W tym przypadku również jest widoczna nieidealność obwodu będąca skutkiem nieidealności wzmacniaczy operacyjnych.



Rys. 9. Przebiegi czasowe prądów i napięć w obwodzie obciążonym rezystorem nieliniowym  $G_n = 10^{-3}u_n$

#### 4. ZŁOŻONE UKŁADY DYNAMICZNE

Jeśli przez  $x[n]$  oznaczymy wektor zmiennych stanu, przez  $e[m1]$  wymuszenie a przez  $y[m2]$  odpowiedź układu nieliniowego gdzie  $n$ ,  $m1$  oraz  $m2$  oznaczają odpowiedni wymiar wektora to nieliniowy układ dynamiczny będziemy opisywać w postaci normalnej [2, 3]

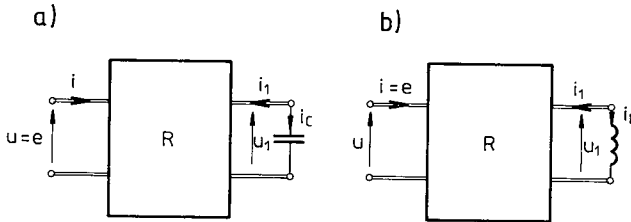
$$\dot{x} = F(x, e); \quad y = G(x, e). \quad (10)$$

Układowi równań (10) przyporządkować można nieliniowy obwód dynamiczny zawierający liniowe pojemności lub indukcyjności oraz nieliniowe rezystancje pasywne bądź aktywne, dwu lub wielokońcówkowe [1, 2, 3]. Dla przykładu jeśli przyjmiemy że zmienne stanu  $x$  reprezentują ładunki kondensatorów  $x = q$ ,  $e$  — wymuszenia napięciowe  $u$ , a wektor odpowiedzi  $y$  prądy w obwodzie  $y = i$  wówczas równanie (10) przekształca się w opis admitancyjny nieliniowego wielowrotnika rezystancyjnego

obciążonego liniowymi pojemnościami (rys. 10a) a opis wielowrotnika rezystancyjnego ma wówczas postać

$$i_1 = -F(Cu_1, u), \quad i = G(Cu_1, u), \quad (11)$$

w której  $C$  oznacza diagonalną macierz odpowiadającą pojemnościom obciążającym wielowrotnik rezystancyjny. Podobnie przyjmując  $x = \psi$  (strumień skojarzony cewek) przy wymuszeniu  $e = i$  i odpowiedziach  $y = u$  otrzymuje się symulację zależności (10) w układzie wielowrotnika rezystancyjnego nieliniowego opisanego równaniami typu impedancyjnego



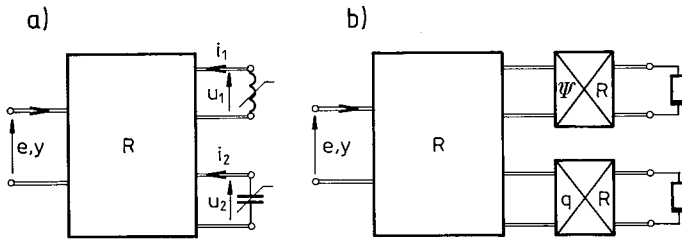
Rys. 10. Nieliniowy model obwodowy odpowiadający równaniu stanu (10) z liniowymi reaktancjami obciążenia; a) obciążenie pojemnościowe, b) obciążenie indukcyjne

$$u_1 = F(-Li_1, i), \quad u = G(-Li_1, i) \quad (12)$$

obciążonego na wrotach wyjściowych indukcyjnościami  $L$  (rys. 10b). W ten sposób realizacja dynamicznego układu sprowadza się do problemu wielowrotnika rezystancyjnego. W ogólnym przypadku realizacja tego wielowrotnika jest problemem stosunkowo złożonym szczególnie jeśli weźmie się pod uwagę, że jedynymi dopuszczalnymi elementami nieliniowymi w praktyce są rezystory nieliniowe dwukońcówkowe a w schemacie realizacyjnym chcemy uniknąć nieliniowych źródeł sterowanych. W wielu wypadkach, szczególnie jeśli zmienne stanu są separowalne, wygodniej jest posłużyć się układem rezystancyjnym obciążonym nieliniowymi pojemnościami i indukcyjnościami. Przy założeniu, że  $x = (q, \psi)$  równania stanu (10) przekształcają się do postaci

$$\begin{aligned} \frac{dx}{dt} &= F\{q(u_c), \psi(i_L), e\} \\ y &= G\{q(u_c), \psi(i_L), e\}, \end{aligned} \quad (13)$$

w której  $q(u_c)$  oraz  $\psi(i_L)$  reprezentują nieliniowe pojemności i indukcyjności,  $dx/dt$  odpowiednio prądy i napięcia natomiast  $y$  oraz  $e$  prądy bądź napięcia odpowiadające bądź odpowiedziom bądź wymuszeniom w obwodzie. Jeśli wielkości  $q(u_c)$  i  $\psi(i_L)$  są separowalne i dadzą się przedstawić w postaci funkcji poszczególnych zmiennych  $q_i, \psi_i$  wówczas zależności (13) odpowiada schemat elektryczny obwodu z rys. 11a w którym wielowrotnik rezystancyjny jest łatwo realizowalny a nieliniowe pojemności i indukcyjności dwukońcówkowe stanowiące obciążenie obwodu realizuje się poprzez nieliniową rezystancję obciążającą odpowiedni typ mutatora (rys. 11b).



Rys. 11. Nieliniowy model obwodowy odpowiadający równaniu stanu (10) z nieliniowym obciążeniem reaktancyjnym; a) obciążenie wrót indukcyjne i pojemnościowe, b) z zastosowaniem mutatorów

Jako przykład takiego podejścia rozważmy realizację układu dynamicznego opisanego następującym równaniem stanu

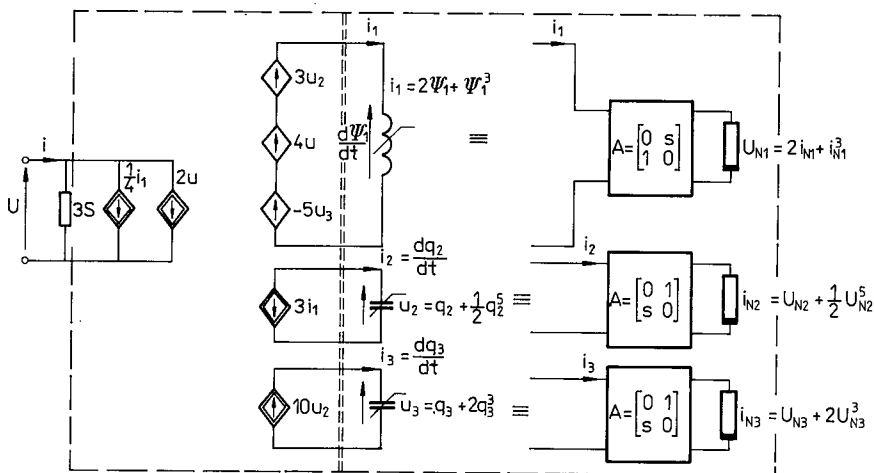
$$\begin{aligned}\dot{x}_1 &= 3(x_2 + 0.5x_2^5) - 5(x_3 + 2x_3^3) + 4u \\ \dot{x}_2 &= -3(2x_1 + x_1^3) \\ \dot{x}_3 &= 5(2x_2 + x_2^5) \\ y &= i = 0.5(x_1 + 0.5x_1^3) + 2(x_3 + 2x_3^3) + 3u\end{aligned}\quad (14)$$

Przyjmując np.  $x_1 = \psi_1$ ,  $x_2 = q_2$ ,  $x_3 = q_3$  oraz  $i_1 = 2\psi_1 + \psi_1^3$ ,  $u_2 = q_2 + 0.5q_2^5$ ,  $u_3 = q_3 + 2q_3^3$  układ równań (14) przekształca się do postaci:

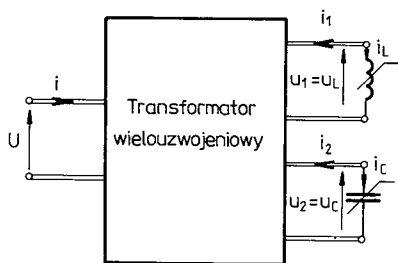
$$\begin{aligned}u_1 &= \frac{d\psi_1}{dt} = 3u_2 - 5u_3 + 4u \\ i_2 &= \frac{dq_2}{dt} = -3i_1 \\ i_3 &= \frac{dq_3}{dt} = 10u_2 \\ i &= 0.25i_1 + 2u_3 + 3u.\end{aligned}\quad (15)$$

Na rys. 12 przedstawiono schemat obwodu nieliniowego opisanego równaniem (15). Część rezystancyjna obwodu jest liniowa a nieliniowości występują jedynie w elementach reaktancyjnych dwukońcówkowych stanowiących obciążenie 4-wrotnika. Elementy reaktancyjne obciążenia realizuje się w postaci odpowiedniego typu mutatora obciążonego rezystorem nieliniowym jak to przedstawiono na rys. 12. Zauważmy, że realizując zależność (14) wg rysunku 10 należałoby użyć wprawdzie liniowych reaktancji ale nieliniowości przeniesione zostałyby do części rezystancyjnej obwodu w postaci sześciu nieliniowych źródeł sterowanych których realizacja jest bardziej złożona niż elementów nieliniowych dwukońcówkowych.

Interesującą klasę dynamicznych układów nieliniowych stanowią pasywne układy bezstratne spełniające warunki wzajemności [2, 3]. Do ich realizacji wystarcza użycie liniowego transformatora wielouzwojeniowego obciążonego na wybranej liczbie wrót nieliniowymi reaktancjami. Schemat ogólny takiego obwodu przedstawia rys. 13.



Rys. 12. Rysunek do przykładu pierwszego



Rys. 13. Struktura kanoniczna nieliniowego obwodu dynamicznego wzajemnego zawierającego dwa rodzaje elementów

Oznaczając kierunki prądów poszczególnych wrót jak na rysunku, transformator wielozwojowy opisuje relacja

$$\begin{bmatrix} u_1 \\ i_2 \\ i \end{bmatrix} = \begin{bmatrix} T'_1 & T'_2 & 0 \\ 0 & T_1 & \\ & & T_2 \end{bmatrix} \begin{bmatrix} u_2 \\ u \\ i_1 \end{bmatrix} \quad (16)$$

Biorąc pod uwagę, że  $V_L = d\psi/dt = u_1$  a  $i_c = dq/dt = -i_2$  z zależności (16) otrzymuje się opis układu równaniami stanu

$$\begin{bmatrix} \frac{d\psi}{dt} \\ \frac{dq}{dt} \end{bmatrix} = \begin{bmatrix} 0 & T'_1 \\ -T_1 & 0 \end{bmatrix} \begin{bmatrix} i_L \\ u_c \end{bmatrix} + \begin{bmatrix} T'_2 \\ 0 \end{bmatrix} u, \quad (17)$$

$$i_2 = [T_2 \quad 0] \begin{bmatrix} i_L \\ u_c \end{bmatrix}. \quad (18)$$

Jeśli wektor stanu oznaczmy przez  $x$  gdzie  $x = [\psi_L, q_c]^T$  natomiast funkcję stanu przez  $g(x)$  gdzie  $g(x) = [i_L(\psi), u_c(q)]^T$  to zakładając że odpowiedź  $y$  stanowią prądy  $i$ ,  $y = i$  odpowiadające wrotom wejściowym a wymuszenia napięcia tych wrót równania stanu obwodu przyjmują postać:

$$\dot{x} = \begin{bmatrix} 0 & T_1^t \\ -T_1 & 0 \end{bmatrix} g(x) + \begin{bmatrix} T_2^t \\ 0 \end{bmatrix} u. \quad (19)$$

Jak z powyższego wynika obwód z rys. 12 pozwala zrealizować układ nieliniowy opisany równaniami stanu o postaci

$$\dot{x} = F(x) + Au, \quad y = G(x). \quad (20)$$

Warunkiem koniecznym i wystarczającym tej realizacji jest istnienie macierzy rzeczywistych  $T_1$  i  $T_2$  oraz funkcji stanu  $g(x)$  tak dobranych aby

$$\begin{aligned} F(x) &= \begin{bmatrix} 0 & T_1^t \\ -T_1 & 0 \end{bmatrix} g(x) \\ A &= \begin{bmatrix} T_2^t \\ 0 \end{bmatrix} \\ G(x) &= [T_2 \quad 0] g(x). \end{aligned} \quad (21)$$

Przy zadanych z góry wektorach  $F(x)$ ,  $G(x)$ ,  $A$  z porównania odpowiednich zależności jest możliwa identyfikacja poszukiwanych elementów macierzy  $T_1$  i  $T_2$  oraz funkcji stanu  $g(x)$ .

Dla zilustrowania sposobu identyfikacji parametrów obwodu z rys. 12 rozpatrzmy przykład liczbowy w którym równania stanu dane są w postaci

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} -5(3x_2 + 0.5x_2^3) \\ 2(x_1 + 2x_1^5) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \\ y &= 0.5x_1 + x_1^5 \end{aligned} \quad (22)$$

Korzystając z równości (21) otrzymuje się układ równań do identyfikacji parametrów przy założeniu, że

$$x = \begin{bmatrix} \psi_L \\ q_c \end{bmatrix}, \quad g(x) = \begin{bmatrix} i_L(\psi_L) \\ u_c(q_c) \end{bmatrix}.$$

Układ ten ma postać

$$\begin{aligned} \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} T_2^t \\ 0 \end{bmatrix} \\ 0.5x_1 + x_1^5 &= [T_2 \quad 0] \begin{bmatrix} i_L(\psi_L) \\ u_c(q_c) \end{bmatrix} \\ \begin{bmatrix} 0 & T_1^t \\ -T_1 & 0 \end{bmatrix} \begin{bmatrix} i_L(\psi_L) \\ u_c(q_c) \end{bmatrix} &= \begin{bmatrix} -5(3x_2 + 0.5x_2^3) \\ 2(x_1 + 2x_1^5) \end{bmatrix}. \end{aligned}$$

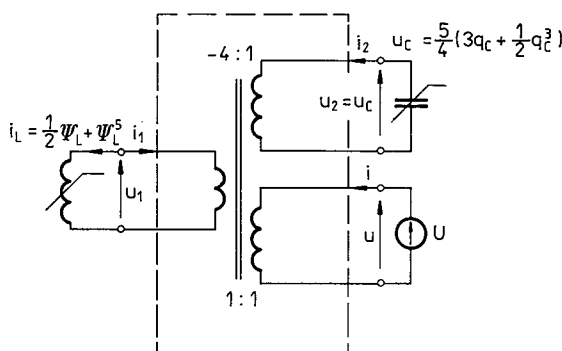
Z rozwiązania powyższego układu równań otrzymuje się kolejno

$$\begin{aligned} T_2 &= 1 \\ i_L(\psi_L) &= 0.5x_1 + x_1^5 \\ T_1 &= -4 \\ u_c(q_c) &= 1.25(3x_2 + 0.5x_2^3), \end{aligned}$$

Stąd pełny opis transformatora 3-uzwojeniowego przyjmuje postać

$$\begin{bmatrix} u_1 \\ i_2 \\ i \end{bmatrix} = \begin{bmatrix} -4 & 1 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ u \\ i_1 \end{bmatrix}$$

Pełny schemat obwodu realizującego układ równań stanu przedstawiono na rys. 14.



Rys. 14. Rysunek do przykładu drugiego

Przykład omówiony powyżej jest szczególnym przypadkiem obwodu nieliniowego wzajemnego który stosunkowo łatwo można było zrealizować. Niestety w ogólnym przypadku sprawdzenie warunków realizowalności takiego obwodu jest problemem skomplikowanym i często trudniejszym niż sam proces syntezy.

## WNIOSKI KOŃCOWE

W pracy dokonano przeglądu metod realizacji elementów i układów dynamicznych przy wykorzystaniu teorii obwodów nieliniowych rezystancyjnych. W przypadku pojedynczych elementów reaktancyjnych przedstawiono zastosowanie czwórników aktywnych (mutatorów). Podano rozwiązania układowe mutatorów klasy konwertera i inwertera impedancyjnego pozwalającego na realizację dowolnego typu elementu dynamicznego. Układy przedstawione w pracy poddano badaniom symulacyjnym uzyskując rezultaty potwierdzające użyteczność prezentowanych rozwiązań w zakresie częstotliwości do około 8 kHz. W badaniach praktycznych uwaga została skupiona na mutatorach klasy R-L ze względu na ich praktyczne znaczenie w realizacji bezindukcyjnej układów. Porównując działanie opracowanego modelu konwertera z

tradycyjnym rozwiązaniem GIC-u (tam gdzie to jest możliwe) na wzmacniaczach operacyjnych stwierdzono szerszy zakres częstotliwości działania proponowanego rozwiązania.

W przypadku złożonych układów dynamicznych obejmujących kilka rodzajów elementów rozważono związek między opisem w przestrzeni stanu a ich reprezentacją obwodową. Dokonano przeglądu metod transformacji jednego rodzaju opisu w drugi ze szczególnym uwzględnieniem opisu wielowrotnika rezystancyjnego. Podano przykłady liczbowe ilustrujące możliwości i ograniczenia takich transformacji.

## BIBLIOGRAFIA

1. L. O. Chua: *Device modeling via basic nonlinear circuit elements*, IEEE Trans. CAS, 1980, vol. 27, No 11, ss. 1014—1044
2. T. H. Nguyen, B. D. O. Anderson: *Analysis and synthesis of nonlinear reciprocal networks containing two elements types and transformers*. IEEE Trans. CAS, 1980, vol. 27, No 4, ss. 263—276
3. L. O. Chua: *Dynamic nonlinear networks: state of the art*. IEEE Trans. CAS, 1980, vol. 27, No 11, ss. 1059—1088
4. A. Cichocki: *Synteza układów nieliniowych przy użyciu wzmacniaczy operacyjnych i elementów sterowanych*. Zeszyty Naukowe Elektryka, Politechnika Warszawska, 1982, No 67
5. S. Osowski: *Novel realization of converter type mutator*, Electronics Letters, 1986, vol. 22, No 8, ss. 424—425
6. A. Cichocki, S. Osowski, S. Filipowicz: *Unified active structure for arbitrary two-port network realization*. Electronics Letters, 1977, vol. 13, No 11, ss. 328—329

S. OSOWSKI

## NONLINEAR DYNAMIC ELEMENTS IN THE ASPECT OF PRACTICAL REALIZATION

### Summary

The paper presents the methods of realization of elements and complex dynamic networks using the theory of resistive nonlinear circuits. The chosen solutions of mutators co-operating with resistive elements to realize reactive elements are analysed. A review of transformations from the state space description to circuits corresponding to this description is given.



# Synthesis of time-suboptimal control system using a reduction of the order of initial model

JANUSZ HALAWA

*Instytut Cybernetyki Technicznej, Politechnika Wrocławska*

*The paper received 1988.11.10*

*Authorized 1989.03.30*

Considerable computing difficulties are encountered when determining the control operations for time-optimal systems in the case of a model described by linear differential equations with constant coefficients when the characteristic roots are real, negative and distinct. The determination of this control makes it necessary to solve the sets of transcendental equations. The problem of substituting an optimal system by a suboptimal one in which the control is time-optimal but for a lower-order system determined from the initial one by means of reduction.

Considerable computing difficulties are encountered when determining the control operations for time-optimal systems in the case of a model described by linear differential equations with constant coefficients when the characteristic roots are real, negative and distinct. The determination of this control makes it necessary to solve the sets of transcendental equations. The paper discusses the problem of substituting an optimal system by a suboptimal one in which the control is time-optimal but for a lower-order system determined from the initial one by means of reduction.

A large number of papers have been written on the problem of time-optimal control. This problem has also been studied by Polish scientists [8, 9, 14].

This paper discusses the problem of synthesis of a time-suboptimal system. The time-suboptimal control system is understood here as a control system in which the input function corresponds to optimal control for a lower-order system obtained by reduction of the initial model. The transfer function describing the simplified model (of lower order than the transfer function describing the initial model) is determined by means of the continued fractions procedure. To do so, the initial transfer function is expanded into a continued fraction and replaced by the approximant of this expansion. This paper deals with a case of synthesis for a time-optimal system described by the transfer function having the form

$$G(s) = \frac{y(s)}{x(s)} = \frac{c_n s^n + c_{n-1} s^{n-1} + \dots + c_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0} \quad (1)$$

Substituting

$$\begin{aligned}
 b_0 &= c_n \\
 b_1 &= c_{n-1} - b_0 a_{n-1} \\
 b_2 &= c_{n-2} - b_0 a_{n-2} - b_1 a_{n-1} \\
 &\dots\dots\dots \\
 b_{n-1} &= c_k - \sum_{i=0}^{n-k-l} b_i a_{i+k}
 \end{aligned} \tag{2}$$

it is possible to describe the model (1) by the differential equation in the following form

$$\dot{\bar{x}} = A\bar{x} + Bu, \tag{3}$$

where

$$\bar{x} = [x_1, x_2, \dots, x_n]^T, \tag{4}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_0 & -a_1 & -a_2 & -a_3 & \dots & -a_{n-1} \end{bmatrix} \tag{5}$$

and

$$B = [b_1, b_2, \dots, b_n]^T. \tag{6}$$

The output function is described by the formula

$$y = x_1 + b_0 u. \tag{7}$$

Let us assume that the roots  $r_1, r_2, \dots, r_n$  of denominator of transfer function describing the initial model and the roots of denominator of simplified transfer function are real, negative and distinct, and that the control function is constrained

$$|u(t)| \leq U. \tag{5}$$

The control operation is to transfer the system from the initial state

$$\bar{x}(0) = [0, 0, \dots, 0]^T \tag{6}$$

in the shortest time possible to the final state

$$\bar{x}(t_n) = [x_1, 0, 0, \dots, 0]^T. \tag{7}$$

The solution of the equation can be presented as follows

$$e^{-At} P^{-1} \bar{x} = \int_0^t e^{-A\tau} P^{-1} B u \, d\tau, \tag{8}$$



a sub-optimal control system, the control of which is the optimal one but for the lower-order determined from the initial system by means of reduction. This control has the number of switching points equal to the order of the simplified transfer function. Thus, this control is easier for practical application. The method of reduction used depends on the particular case concerned. For example, the method of rejecting the less-significant characteristic roots may not be applied when they are close to one another. The frequency methods [2, 6, 11, 12] are very time-consuming. The parameters of the simplified model depend on the pulsation range assumed  $[0, \omega_g]$ . From among these models we then select a model that provides the best approximation for the frequency characteristic of the initial model [6]. Frequency methods [2, 12] often do not permit for satisfactory results. One of the better methods of reduction is the procedure of continued fractions [4, 5, 7], introduced by Chen and Shieh (1968). According to the procedure the initial transfer function is expanded into continued fraction of the form

$$G(s) = \frac{1}{t_0 + \frac{s}{t_1 + \frac{s}{t_2 + \dots}}} = \frac{1}{t_0} + \frac{s}{t_1} + \frac{s}{t_2} + \dots$$

The simplified transfer function is equal to one of the approximants of this expansion. This procedure is used in this paper to reduce the transfer function of initial model. Now, let us consider an example of synthesis of a time-suboptimal system.

### Example

The transfer function of the initial model is

$$G(s) = \frac{y(s)}{x(s)} = \frac{V_1(s)}{V_0(s)} \quad (14)$$

where

$$V_1(s) = 699.83903$$

and

$$V_0(s) = s^5 + 31.89957s^4 + 285.55958s^3 + 526.60325s^2 + 341.92749s + 69.983903.$$

The roots of  $V_0(s)$  polynomial are:  $r_1 = -0.4$ ,  $r_2 = -0.8$ ,  $r_3 = -1$ ,  $r_4 = -13.5$ ,  $r_5 = -16.2$ . We determine the transfer function of the simplified model, the denominator of this function being a third-order polynomial. When expanding the transfer function (14) into a continued fraction of the form

$$G(s) = \frac{1}{t_0} + \sum_{i=1}^k \frac{s}{t_i}, \quad k = 2n-1$$

by the algorithm given in [3] i.e.

Let us  $\deg V_1(s) = p$  and  $\deg V_0(s) = n$   $p \leq n$  for  $i = 1, 2, 3, \dots$

$$t_0 = \frac{V_0(0)}{V_1(0)}$$

$$V_{i+1}(s) = \frac{1}{s} [V_{i-1}(s) - t_{i-1} V_i(s)].$$

$$t_i = \frac{V_i(0)}{V_{i+1}(0)},$$

we obtain two approximants with third-order denominators, namely:

$$G_{33}(s) = \frac{y_{33}(s)}{x(s)} = \frac{-0.003322 s^3 + 0.043322 s^2 - 0.433233 s + 3.198630}{s^3 + 2.199518 s^2 + 1.519466 s + 0.319863} \quad (15)$$

with the denominator roots  $r_1 = -0.999293$ ,  $r_2 = -0.800226$ ,  $r_3 = -0.399999$  and

$$G_{23}(s) = \frac{y_{23}(s)}{x(s)} = \frac{0.036029 s^2 - 0.422419 s + 3.186610}{s^3 + 2.195024 s^2 + 1.514673 s + 0.318661} \quad (16)$$

with the denominator roots  $r_1 = -0.991631$ ,  $r_2 = -0.803411$ ,  $r_3 = -0.399982$ .

The simplified transfer functions satisfy the assumptions having been specified previously because the roots of their denominators are real, negative and distinct. Table 1 summarizes the values of responses for models described by the transfer functions (14), (15) and (16) in the case of the step function  $x(t) = 1(t)$  applied. The functions  $y(t)$ ,  $y_{33}(t)$  and  $y_{23}(t)$  are illustrated in Fig. 1. These graphs in Fig. 1 overlap each other. The gain-phase characteristics of the systems described by the transfer functions (14), (15) and (16) are shown in Fig. 2. Like in the previous case, these characteristics overlap each other in Fig. 2. In the cases of the simplified models considered in this paper the task of the time-optimal control is to transfer the system from the initial state  $[0, 0, 0]^T$  to the final state  $[1, 0, 0]^T$ . Let us assume that  $U = 1$  and the control value in the first interval is positive. The set of transcendental equations (13) for the case of the simplified model described by the transfer function (15) is:

$$\begin{aligned} 2e^{0.4000t'_1} - 2e^{0.4000t'_2} + 0.905338 e^{0.4000t'_3} - 1 &= 0 \\ 2e^{0.8002t'_1} - 2e^{0.8002t'_2} + 0.910521 e^{0.8002t'_3} - 1 &= 0 \\ 2e^{0.9993t'_1} - 2e^{0.9993t'_2} + 0.913039 e^{0.9993t'_3} - 1 &= 0. \end{aligned} \quad (17)$$

Tablica 1

The time response of the initial system (14) and of the simplified systems (15) and (16)

| $t$         | 0                     | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 12   | 14   | 20   |
|-------------|-----------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| $y(t)$      | 0                     | 0.22 | 1.31 | 2.94 | 4.61 | 6.05 | 7.19 | 8.04 | 8.65 | 9.08 | 9.37 | 9.71 | 9.87 | 9.98 |
| $y_{33}(t)$ | $-3.22 \cdot 10^{-5}$ | 0.22 | 1.31 | 2.94 | 4.61 | 6.05 | 7.19 | 8.04 | 8.65 | 9.08 | 9.37 | 9.71 | 9.87 | 9.98 |
| $y_{23}(t)$ | 0                     | 0.22 | 1.31 | 2.94 | 4.61 | 6.05 | 7.19 | 8.04 | 8.65 | 9.08 | 9.37 | 9.71 | 9.87 | 9.98 |

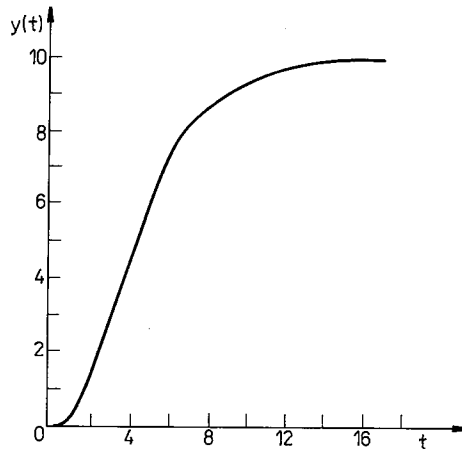


Fig. 1. The time responses of the initial system (14) and of the simplified systems (15) and (16)

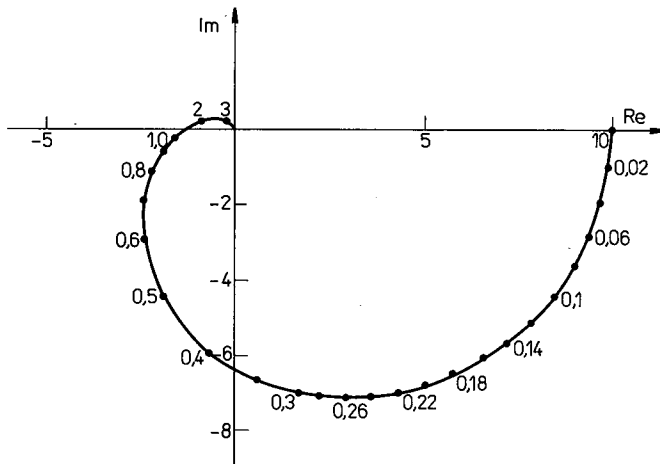


Fig. 2. The frequency responses of the initial system (14) and of the simplified systems (15) and (16)

The solution of this set gives the switching times:

$$t'_1 = 1.0706, \quad t'_2 = 1.9220, \quad t'_3 = 2.2658.$$

In case of the simplified model described by the transfer function (16), the set of transcendental equations is

$$\begin{aligned} 2e^{0.4000t'_1} - 2e^{0.4000t'_2} + 0.905201e^{0.4000t'_3} - 1 &= 0 \\ 2e^{0.8034t'_1} - 2e^{0.8034t'_2} + 0.910218e^{0.8034t'_3} - 1 &= 0 \\ 2e^{0.9916t'_1} - 2e^{0.9916t'_2} + 0.912478e^{0.9916t'_3} - 1 &= 0. \end{aligned} \quad (18)$$

The solution of this set gives the switching times

$$t''_1 = 1.0716, \quad t''_2 = 1.9220, \quad t''_3 = 2.2697.$$

The determination of the above switching times for the simplified model makes it possible to find the initial approximations of the solution for the set of transcendental equations from which the switching times for the initial model described by the transfer function (14) are calculated. In the case of the initial model the task of the time-optimal-control is to transfer the system from the initial state  $[0, 0, 0, 0]^T$  to the final state  $[1, 0, 0, 0]^T$ . The set of the transcendental equations is as follows

$$\begin{aligned} e^{0.4t_1} - e^{0.4t_2} + e^{0.4t_3} - e^{0.4t_4} + 0.449923e^{0.4t_5} - 0.5 &= 0 \\ e^{0.8t_1} - e^{0.8t_2} + e^{0.8t_3} - e^{0.8t_4} + 0.449996e^{0.8t_5} - 0.5 &= 0 \\ e^{t_1} - e^{t_2} + e^{t_3} - e^{t_4} + 0.450046e^{t_5} - 0.5 &= 0 \\ e^{13.5t_1} - e^{13.5t_2} + e^{13.5t_3} - e^{13.5t_4} + 0.456572e^{13.5t_5} - 0.5 &= 0 \\ e^{16.2t_1} - e^{16.2t_2} + e^{16.2t_3} - e^{16.2t_4} + 0.453704e^{16.2t_5} - 0.5 &= 0. \end{aligned} \quad (19)$$

The initial approximations  $t_1, t_2, t_3$  have been assumed as the switching times determined for the simplified model. On the other hand, the remaining approximations are proximate to  $t_3$  and in our case, they are  $t_4 = 2.28, t_5 = 2.33$ . The solution of the set of equations (13) gives the switching times:  $t_1 = 1.0717, t_2 = 1.923, t_3 = 2.270, t_4 = 2.3$  and  $t_5 = 2.35$ . A modified Powell method has been used to resolve this set of equations. The method used to solve the set of non-linear equations requires a fairly precise determination of the initial approximations. If the time-optimal control  $u_{33}(t)$  for the simplified model described by the transfer function (15) is assumed to be the input

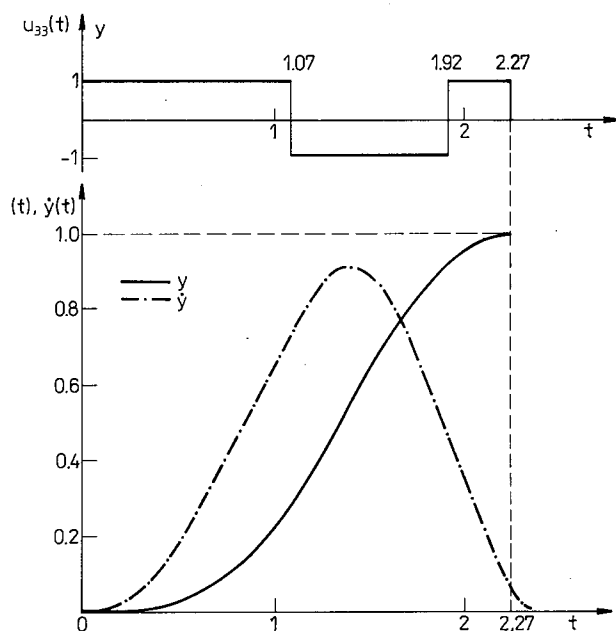


Fig. 3. The time responses of the initial system (14) for the time-optimal control function of the simplified system (15)

function of the initial system described by the transfer function (14), we obtain  $y_{u_{33}}(t'_3) = y_{u_{33}}(2.2658) = 0.9951$ . The relative error in reaching the present value  $l$  is, in this case, equal  $\delta_{33} = -0.49\%$ . For the  $u_{23}(t)$  control, we obtain  $y_{u_{33}}(t''_3) = y_{23}(2.2697) = 0.9972$  thus  $\delta_{23} = -0.28\%$ . These accuracies seem to be acceptable in some practical cases. No successful results have been obtained when the assumption was made that the control value in the first interval is negative.

## CONCLUSIONS

As shown for the example under consideration it is possible, in some cases, to make a synthesis of a time-suboptimal control according to the suggested procedure. The response of the initial system to time-optimal control for this system are well approximated by its responses to suboptimal controls, i.e. the optimal controls for simplified systems. Since a good approximation of the responses of the initial system with these of the simplified system is required, the suboptimal controls determined for the simplified systems of the third or higher orders should be considered. The switching times of suboptimal controls may be used as initial approximations in optimization methods used to find the time-optimal controls for initial model. The responses of the stable simplified systems determined by means of the continued fraction procedure provide, generally, a good approximation for the response of a stable initial system. When using the procedure of continued fractions one should take into account the fact that there may be some complex roots of approximant denominators though all the roots of the denominator of the transfer function describing the initial model are real. There may be approximants with unstable denominators through the denominator of the transfer function for the initial model is stable. Such approximants cannot be used.

## REFERENCES

1. M. Athans, P. L. Falb: „Sterowanie optymalne. Wstęp do teorii i jej zastosowania”, WNT, Warszawa, 1969
2. Y. Bistritz, G. Langholz: *Model Reduction by Chebyshev Polynomial Techniques*. IEEE Trans. on Autom. Control, AC-24, no. 5, 1979
3. A. Burtheel: *Division algorithms for continued fractions and the Pade table*. Report TW41, August 1978, Applied Mathematics and Programming Division, Katholieke Universitet Leuven (Belgium)
4. C. F. Chen: *Model reduction of multivariable control systems by means of matrix continued fractions*. Int. J. Control, 20, no. 2, 1974.
5. C. E. Chen, L. S. Shieh: *A novel approach to linear model simplification*. Int. J. Control, 8, no. 6, 1968.
6. J. Halawa: *Comments on A new frequency domain techniques for the simplification of linear dynamic systems and Method for frequency domain simplification of transfer functions*. Int. J. Control, 41, no. 1, 1985
7. J. Halawa, A. Trzmielak-Stanisławska: *Properties of Simplified Transfer Functions as Acquired with a Method of Continued Fractions*. Archiwum Automatyki i Telemekhaniki, tom XXXI, zeszyt 1—2, 1986
8. W. Hejmo: *On the Sensitivity of a Time-Optimal Positional Control*. IEEE Transactions on Automatic Control, vol. AC-28, no. 5, 1983

9. W. Hejmo, P. Kulczycki: *O rozwiązaniu problemu czasowo-optimalnego sterowania pozycyjnego*. Archiwum Automatyki i Telemechaniki, tom. XXXI, zeszyt 1—2, 1986
10. T. Kaczorek: *Macierze w automatyce i elektrotechnice*, WNT, Warszawa 1984
11. S. V. Rao, S. S. Lamba: *A new frequency domain technique for the simplification of linear dynamic systems*, Int. J. Control, 20, no. 1, 1974
12. A. S. S. R. Reddy: *A method for frequency domain simplification of transfer functions*, Int. J. Control, 23, no. 3, 1976
13. H. R. Sirisena: *A gradient method for computing optimal bang-bang controls*, Int. J. Control, 19, no. 2, 1974
14. A. Wierzbicki: *Modele i wrażliwość układów sterowania*, WNT, Warszawa, 1977

JANUSZ HALAWA

## SYNTEZA CZASO-OPTYMALNEGO UKŁADU STEROWANIA Z WYKORZYSTANIEM REDUKCJI RZĘDU MODELU WYJŚCIOWEGO

### Streszczenie

Wyznaczanie czasów przełączeń sterowania czasowo-optimalnego w przypadku modeli opisanych stabilnymi liniowymi równaniami różniczkowymi wyższych rzędów ze stałymi współczynnikami, których pierwiastki charakterystyczne są rzeczywiste, ujemne i różne, nastręcza duże trudności obliczeniowe ze względu na konieczność rozwiązywania układów równań transcendentnych. W niniejszej pracy rozważana jest możliwość zastąpienia układu optymalnego układem suboptymalnym, w którym sterowanie jest sterowaniem optymalnym lecz dla układu niższego rzędu otrzymanego z układu wyjściowego drogą redukcji.



# Prosta metoda projektowania dwuwymiarowych filtrów cyfrowych

WIESŁAW MARSZAŁEK

*Instytut Elektrotechniki, Wyższa Szkoła Inżynierska w Opolu*

*Otrzymano 1988.04.13*

*Autoryzowano do druku 1989.04.25*

W pracy przedstawiono metodę projektowania dwuwymiarowych filtrów cyfrowych w której wykorzystano transformację częstotliwości  $z = z'_1 z'_2$ ;  $z'_i = (z_i + a)/(az_i + 1)$ ,  $|a| < 1$ ,  $i = 1, 2$ . Transformacja ta prowadzi do stabilnych filtrów dwuwymiarowych o charakterystykach amplitudowych mających kształt kołowy. Przedstawiono dwa przykłady projektowania dwuwymiarowych filtrów dolno- i górnoprzepustowych.

## 1. WPROWADZENIE

Rozwój teorii sygnałów dwuwymiarowych i zapotrzebowania praktyki na efektywne metody i środki cyfrowego przetwarzania tych sygnałów spowodował duże zainteresowanie wielu autorów tą dziedziną badań. Metody i algorytmy teorii układów i sygnałów dwuwymiarowych oparte są w wielu przypadkach na odpowiednich uogólnieniach teorii układów jednowymiarowych [1–8].

Cyfrowe przetwarzanie sygnałów dwuwymiarowych znajdujące zastosowanie w analizie obrazów wizyjnych do celów robotyki, medycyny, geodezji, kartografii, mechaniki i in. charakteryzuje się dużą złożonością obliczeniową, wymaganiami dotyczącymi szybkości i objętości przetwarzania danych liczbowych. Na przykład, znany obecnie i dostępny w kraju sprzęt mikrokomputerowy klasy IBM PC pozwala na rejestrację w czasie rzeczywistym obrazów o rozdzielczości  $512 \times 512$  i 256 poziomach jasności.

Efektywność i szybkość przetwarzania sygnałów dwuwymiarowych zależy m.in. od celu przetwarzania, przyjętej metody realizacji i klasy stosowanego sprzętu cyfrowego. Duże różnice w złożoności obliczeniowej i szybkości uzyskiwania wyników zaobserwowano na przykład w realizacji zadań filtracji zakłóceń obrazów stacjonarnych przedstawionych w pracy [9].

W pracy tej podano metodę projektowania filtrów cyfrowych dwuwymiarowych wynikającą z odpowiedniej transformacji częstotliwości. Zasadniczą istotę metody stanowi projekt filtru-prototypu jednowymiarowego oraz przekształcenie jego transmitancji  $H(z)$  w transmitancję dwuwymiarową  $F(z_1, z_2)$ . Przekształcenie to cechuje duża prostota pozwalająca na uzyskanie filtru o charakterystyce amplitudowej zbliżonej w kształcie do koła.

## 2. METODA WYZNACZANIA TRANSMITANCJI DWUWYMIAROWEJ

Niech  $H(z)$  będzie transmitancją filtru jednowymiarowego dla której zastosowano przekształcenie częstotliwości o postaci [10]

$$z = z'_1 z'_2, \quad z'_i = \frac{z_i + a}{az_i + 1}, \quad |a| < 1, \quad i = 1, 2. \quad (1)$$

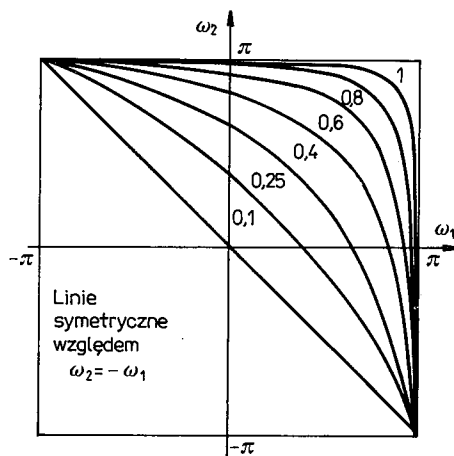
Uwzględniając zależność

$$z_i = \exp(j\omega_i), \quad \omega_i \in (-\pi, \pi) \quad (2)$$

z równania (1) otrzymamy związek określający transformację częstotliwości jednowymiarowej w dwuwymiarową o następującej postaci

$$\omega = \omega_1 + \omega_2 - 2 \left( \operatorname{tg}^{-1} \frac{a \sin \omega_1}{a \cos \omega_1 + 1} + \operatorname{tg}^{-1} \frac{a \sin \omega_2}{a \cos \omega_2 + 1} \right). \quad (3)$$

Transformacja (1) charakteryzuje się symetrią względem prostych  $\omega_2 = \pm \omega_1$ , a linie jednakowych wartości  $\omega$  mają kształt zbliżony do kołowego dla  $(0, 0) \leq (\omega_1, \omega_2) \leq (\pi, \pi)$  oraz  $(-\pi, -\pi) \leq (\omega_1, \omega_2) \leq (0, 0)$ . Przykładowe linie jednakowych wartości  $\omega$  (wzór (3)) dla  $a = 0,6$  przedstawiono na rys. 1.



Rys. 1. Linie stałych wartości  $\omega/\pi$  (wzór (3))

Jeżeli w miejsce transformacji (1) uwzględnimy zależność

$$z = z'_1 z'_2, \quad z'_1 = \frac{z_1 + a}{az_1 + 1}, \quad z'_2 = \frac{z_2^{-1} + a}{az_2^{-1} + 1} \quad (4)$$

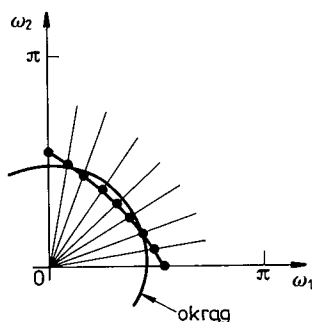
wówczas uzyskamy transformację częstotliwości jednowymiarowej w dwuwymiarową, dla której linie jednakowych wartości  $\omega$  mają kształt kołowy w II i IV części układu współrzędnych (linie te powstają przez obrót linii z rys. 1 o  $90^\circ$ ).

Uwzględniając powyższy fakt symetrii kołową dla transmitancji dwuwymiarowej w całym obszarze  $-\pi \leq \omega_i \leq \pi, i = 1, 2$  uzyskamy przez kaskadowe połączenie filtru o transmitancji  $H\left(\frac{z_1 + a}{az_1 + 1} \cdot \frac{z_2 + a}{az_2 + 1}\right)$  z filtrem o transmitancji  $H\left(\frac{z_1 + a}{az_1 + 1} \cdot \frac{z_2^{-1} + a}{az_2^{-1} + 1}\right)$ .

Określenie wartości rzeczywistej  $a$  w rozważanej transformacji dokonuje się dla żądanej wartości promienia  $r = \sqrt{\omega_1^2 + \omega_2^2}$  wyznaczającego przybliżenie częstotliwości dwuwymiarowej. Wyznaczenie wartości  $a$  polega wtedy na minimalizacji błędu

$$e = \min_{|a| < 1} \frac{\sum_{i=1}^N \left| r - \sqrt{\omega_{1i}^2 + \omega_{2i}^2} \right|}{Nr} \quad (5)$$

gdzie  $N$  oznacza liczbę punktów I części płaszczyzny  $(\omega_1, \omega_2)$  o równomiernym rozkładzie kątowym (rys. 2). Przykładowe wartości  $a$  oraz odpowiadające im wartości błędu (5) przedstawiono w tabeli 1.

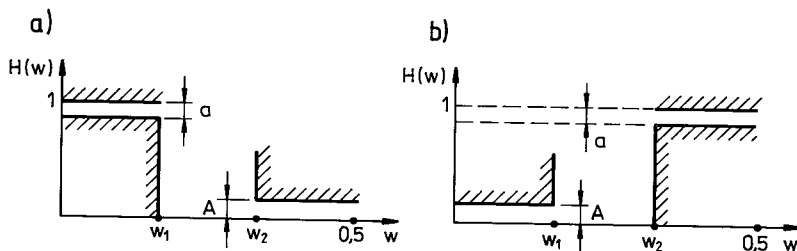


Rys. 2. Przybliżenie kołowe linii stałej wartości  $\omega$

### 3. PRZYKŁADY NUMERYCZNE

W wyznaczaniu transmitancji dwuwymiarowego filtru cyfrowego istotną rolę odgrywa projektowanie jego prototypu jednowymiarowego, którego transmitancja jest następnie przekształcana do postaci dwuwymiarowej za pomocą transformacji (1). Projektowania filtru jednowymiarowego można dokonać w oparciu o jedną z wielu znanych i dostępnych w tym zakresie metod i programów komputerowych. W opracowanym przez autora programie wyznaczania transmitancji filtru dwuwymiarowego skorzystano na etapie projektowania prototypu jednowymiarowego z części algorytmu i programu z monografii [11]. Danymi wejściowymi do programu określającego transmitancję prototypu jednowymiarowego są: typ filtru (dolno-, górnoprzepustowy), odpowiednie częstotliwości charakterystyki amplitudowej filtru (rys. 3), rodzaj aproksymacji (Butterwortha, Czebyszewa), dopuszczalne odchylenie i minimalne tłumienie charakterystyki w odpowiednich pasmach oraz rodzaj struktury

filtru (istnieje możliwość wyboru jednej z dwóch struktur przedstawionych w pracy [11]). Przyjęto, że filtr jednowymiarowy będzie kaskadowym połączeniem struktur o transmitancjach rzędu 1 lub 2 uszeregowanych według rosnących lub malejących dobroci poszczególnych struktur. Szczegółowy opis własności struktur, ich schematy i analiza matematyczna przedstawione są w [11].



Rys. 3. Podstawowe parametry charakterystyk amplitudowych dolnoprzepustowych (rys. a) i górnoprzepustowych (rys. b) w układzie ze znormalizowaną częstotliwością  $\omega T/2\pi$ ;  $-20 \lg(1-a)$  [dB] oznacza maksymalne odchylenie charakterystyki w paśmie przepustowym;  $-20 \lg A$  [dB] jest minimalnym tłumieniem w paśmie zaporowym

Po określeniu transmitancji filtru jednowymiarowego dokonane zostaje jej przekształcenie w odpowiednią transmitancję dwuwymiarową.

### Przykład 1

Rozważmy zadanie wyznaczenia transmitancji dwuwymiarowego filtru cyfrowego dolnoprzepustowego o kołowej charakterystyce amplitudowej z częstotliwością zakresu przepuszczania  $r = 0,55\pi$ . Zakładamy, że charakterystyka filtru jednowymiarowego będzie mieć aproksymację typu Butterwortha z dopuszczalnym odchyleniem 3 dB i minimalnym tłumieniem w paśmie zaporowym 30 dB. Zakłada się ponadto, że filtr będzie kaskadowym połączeniem struktur typu 1 [11, str. 256].

Tablica 1 [10]

Dobór wartości  $a$  w przekształceniu częstotliwości zgodnie z wzorem (1)

| $r/\pi$ | $a$  | $e$ % |
|---------|------|-------|
| 0,50    | 0,9  | 12    |
| 0,60    | 0,9  | 7,6   |
| 0,70    | 0,9  | 1,7   |
| 0,75    | 0,61 | 1,2   |
| 0,80    | 0,49 | 1,3   |
| 0,85    | 0,42 | 1,4   |
| 0,90    | 0,39 | 1,4   |
| 0,95    | 0,36 | 1,4   |
| 1,00    | 0,35 | 1,5   |

Uwzględniając wzór (3) dla danych  $(\omega_1, \omega_2) = (0, r)$  (lub  $(r, 0)$ ) otrzymujemy odpowiadającą częstotliwości dwuwymiarowej  $0,55\pi$  wartość częstotliwości filtru jednowymiarowego ( $a = 0,9$  — tabela 1)

$$\omega = r - 2\operatorname{tg}^{-1} \frac{a \sin r}{a \cos r + 1} = 0,55\pi - 2\operatorname{tg}^{-1} \frac{0,9 \sin 0,55}{0,9 \cos 0,55 + 1} = 0,123\pi \text{ rad/s.} \quad (6)$$

Wyniki obliczeń numerycznych współczynników liczbowych wielomianów transmitancji filtru przy założonym tłumieniu 30 dB dla  $\omega > 0,188\pi$  rad/s przedstawiono w tabeli 2. Z wyników tych wynika, że poszukiwany filtr ma rząd transmitancji równy 9 a poszczególne struktury (I lub II rzędu) są uszeregowane według wzrastającej dobroci. W przypadku struktury II rzędu jej dobroć określona jest wzorem

$$Q = \Theta r / (1 - r^2), \quad \Theta = \operatorname{tg}^{-1} \sqrt{4b_2/b_1^2 - 1}, \quad r = \sqrt{b_2} \quad (7)$$

Tabela 2

Wyniki obliczeń do przykładu 1

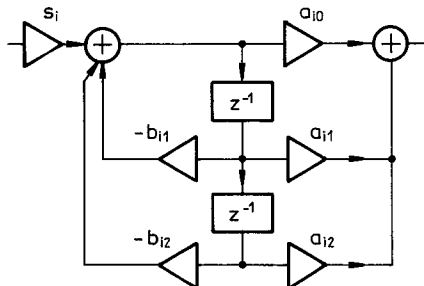
| i<br>Nr<br>struktury | 1       | 2       | 3       | 4       | 5       |
|----------------------|---------|---------|---------|---------|---------|
| $a_{i0}$             | 1,0000  | 1,0000  | 1,0000  | 1,0000  | 1,0000  |
| $a_{i1}$             | 2,0000  | 2,0000  | 2,0000  | 1,0000  | 2,0000  |
| $a_{i2}$             | 1,0000  | 1,0000  | 1,0000  | 0,0000  | 1,0000  |
| $b_{i0}$             | 1,0000  | 1,0000  | 1,0000  | 1,0000  | 1,0000  |
| $b_{i1}$             | -1,7794 | -1,8141 | -1,8700 | -0,8838 | -1,9434 |
| $b_{i2}$             | 0,7930  | 0,8279  | 0,8842  | 0,0000  | 0,9582  |
| $s_i$                | 0,0148  | 0,0195  | 0,0185  | 0,0579  | 0,0196  |
| $Q_i$                | 0,1820  | 0,4200  | 0,8677  | 0,8838  | 2,8426  |

Oznaczenia:

$a_{ik}, b_{ik}$  — współczynnik licznika i mianownika transmisji struktur I lub II rzędu,  $i = 1, \dots, 5$ ;  $k = 0, 1, 2$ .

$s_i, Q_i$  — odpowiednio: współczynnik wzmocnienia i dobroci struktur.

$$H(z) = \prod_{i=1}^5 s_i \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{b_0 + b_1 z^{-1} + b_2 z^{-2}} \quad (\text{transmitacja}).$$



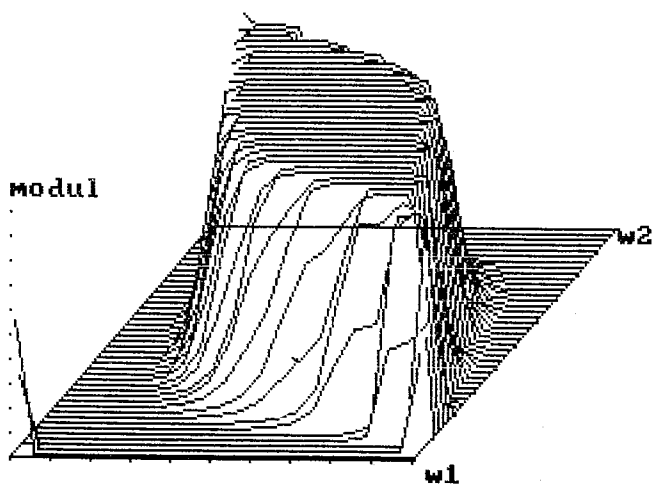
Rys. 4. Elementarna struktura filtru (typ 1 [11])

a dla struktury I rzędu mamy

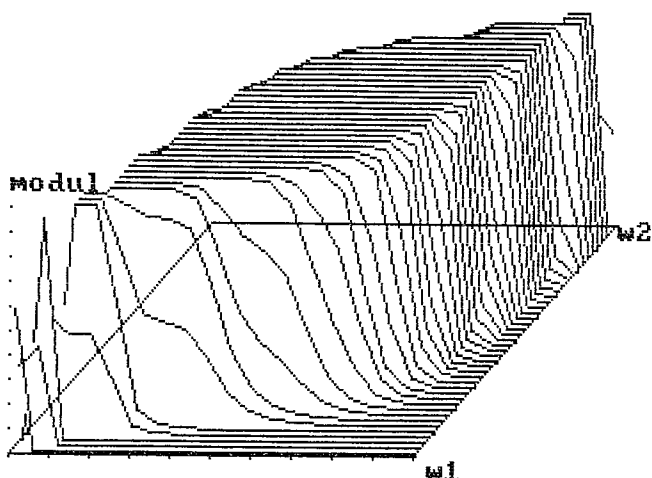
$$Q = |b_1|. \quad (8)$$

Ogólna postać struktury elementarnej przedstawiona jest na rys. 4.

Mając obliczone współczynniki transmitancji filtra jednowymiarowego można łatwo (wzór (1)) obliczyć odpowiednie współczynniki transmitancji dwuwymiarowej zmiennych  $z_1$  i  $z_2$ . Rys. 5 przedstawia charakterystykę amplitudową filtra dwuwymiarowego otrzymaną dla rozważanych wyżej danych. W celu uzyskania charakterystyki symetrycznej w obszarze  $-\pi \leq \omega_1, \omega_2 \leq \pi$  należy dodatkowo włączyć filtr o charakterystyce przedstawionej na rys. 6 (patrz rozdz. 2). Należy dodać, że niewielkie fragmenty pasma przepustowego widoczne na rys. 5 dla wartości  $\omega_1$  i  $\omega_2$  bliskich  $-\pi$  można usunąć stosując filtry jednowymiarowe o odpowiednich zakresach częstotliwości [10].



Rys. 5. Charakterystyka amplitudowa filtra dolnoprzepustowego z przykładu 1

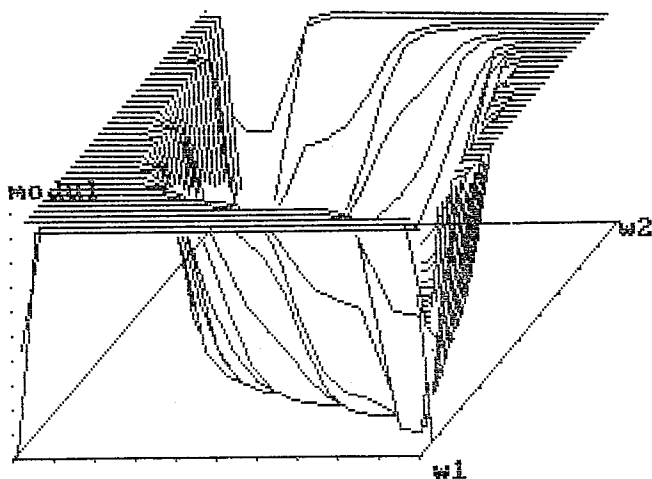


Rys. 6. Charakterystyka z rys. 5 obrócona o  $90^\circ$

## Przykład 2

Przedstawioną szczegółowo w przykładzie 1 metodę projektowania filtrów dwuwymiarowych można również stosować w przypadku filtrów górnoprzepustowych. Na rys. 7 przedstawiono charakterystykę amplitudową takiego filtra otrzymaną dla przypadku stosowania prototypu jednowymiarowego o danych:

- minimalne tłumienie 30 dB dla  $\omega < 0,123 \pi$  rad/s,
- maksymalne odchylenie charakterystyki 3 dB dla  $\omega > 0,188 \pi$  rad/s.



Rys. 7. Charakterystyka amplitudowa filtra górnoprzepustowego z przykładu 2

## WNIOSKI

Przedstawiona metoda wyznaczania transmitancji dwuwymiarowego filtra cyfrowego może być stosowana do projektowania filtrów o charakterystykach kołowych dla częstotliwości  $r = \sqrt{\omega_1^2 + \omega_2^2} > 0,5 \pi$ . W zakresie mniejszych częstotliwości uzyskuje się błędy doboru wartości  $a$  (tabela 1) większe od 10%. Spowodowane jest to kształtem linii odwzorowania (3) będących w zakresie małych częstotliwości w przybliżeniu liniami prostymi (rys. 1). Jak wykazano w pracy [10] niedogodność tą można wyeliminować projektując filtr o częstotliwości  $r_1 = m\pi > 0,5 \pi, m \in N$ . Jeżeli jego transmitancję oznaczmy przez  $T(z_1, z_2)$ , wówczas filtr o transmitancji  $T(z_1^m, z_2^m)$  będzie miał żadaną częstotliwość  $r$  z zakresu  $0—0,5 \pi$  [10].

## BIBLIOGRAFIA

- [1] T. Kaczorek: *Two-dimensional linear systems*, Springer-Verlag, Berlin, 1985
- [2] N.K. Bose: *Multidimensional systems theory: progress, directions and open problems*. Reidel, Dordrecht, 1985

- [3] T. Kaczorek: *Straight line reachability of Roesser model*. IEEE Trans., 1987, vol. AC-32, No. 7, pp. 637—639
- [4] T. Kaczorek: *Deadbeat servoproblem for multivariable  $n$ -D systems*, IEEE Trans., 1986, vol. AC-31, No. 4, pp. 360—362
- [5] Z. Trzaska, W. Marszałek: *Teoria sygnałów i układów elektrycznych*. Wyd. WSI Opole, 1987
- [6] W. Marszałek: *Two-dimensional discrete model with varying coefficients*. IEE Proc., Pt. D, 1986, vol. 133, No. 4, pp. 177—183
- [7] W. Marszałek, J. Sadecki: *Dynamic programming for 2-D discrete linear systems*. IEEE Trans., 1988, vol. AC-33, (accepted for publication)
- [8] W. Marszałek: *Dwuwymiarowe dyskretne modele Roessera dla równania Poissona*. Rozprawy Elektrotechniczne, 1987, tom XXXIII, z. 1, str. 3—10
- [9] W. Marszałek, J. Sadecki: *Separowalne dwuwymiarowe modele różnicowe w filtracji zakłóceń obrazów stacjonarnych*. Mater. XI Semin. Podstaw Elektrotechniki i Teorii Obwodów, Gliwice—Ustroń, kwiecień 1988
- [10] L. Harn, B.A. Shenoi: *Design of stable two-dimensional IIR filters using digital spectral transformations*. IEEE Trans., 1986, vol. CAS-33, No. 5, pp. 483—490
- [11] Ł.M. Goldenberg, B.D. Matiuszkin, M.N. Polak: *Cifrovaja obrabotka signalov*, Radio i Sviaz, Moskwa, 1985

W. MARSZAŁEK

## A SIMPLE METHOD OF DESIGNING OF TWO-DIMENSIONAL DIGITAL FILTERS

### Summary

The paper presents a method of designing of two-dimensional digital filters in which spectral transformation of the form  $z = z'_1 z'_2$  where  $z'_i = (z_i + a)/(az_i + 1)$ ,  $|a| < 1$ ,  $i = 1, 2$  is used. This transformation gives stable circular two-dimensional filters. Two numerical examples of designing of low- and high-pass two dimensional filters are quoted.

# Zastosowanie iteracyjnej metody rozkładu na funkcje własne do analizy dielektrycznych struktur mikrofalowych

Michał Mrozowski, Jerzy Mazur

*Instytut Telekomunikacji, Politechnika Gdańska*

TOMASZ JABŁOŃSKI

*IPPT PAN, Warszawa*

*Otrzymano 1989.03.08*

*Autoryzowano do druku 1989.05.28*

## 1. WSTĘP

W konstrukcji nowoczesnych urządzeń elektronicznych przeznaczonych do pracy w zakresie bardzo wysokich częstotliwości zaliczanych do pasma fal milimetrowych powszechnie stosuje się rezonatory i falowody dielektryczne. Rezonatory spełniają funkcję obwodów rezonansowych umożliwiając konstrukcję selektywnych filtrów, wzmacniaczy oraz generatorów o niskim poziomie szumów FM. Z kolei falowody dielektryczne spełniają dwojaką rolę. Przede wszystkim umożliwiają one integrację poszczególnych części obwodu o stałych rozłożonych a ponadto są istotnym elementem konstrukcyjnym szeregu elementów w tym układów dopasowujących, sprzęgaczy i dzielników mocy.

W ostatnich 15 latach opracowano szereg metod analizy struktur dielektrycznych, począwszy od przybliżonych, do których należy np. metoda efektywnej przenikalności elektrycznej [6], do dokładnych czyli pełnofalowych, spośród których można wymienić metody rodzajów sprzężonych [4], dopasowania rodzajów [3], różnic czy elementów skończonych [5, 7]. Wspólną cechą metod pełnofalowych jest konieczność konstrukcji i rozwiązywania dużych układów równań liniowych. Pociąga to za sobą znaczną komplikację programów komputerowych jak również wydłuża czas obliczeń numerycznych. Problemy te dają o sobie znać szczególnie przy analizie zagadnień trójwymiarowych, do których zalicza się badanie rezonatorów. Dla uzmysłowienia sobie skali problemu można przytoczyć wymagania niektórych programów komputerowych dotyczące rozmiaru pamięci, które niejednokrotnie zajmują obszar rzędu kilkudziesięciu MB! Kolejną wadą wielu metod jest warunek bezstratności struktury, co uniemożliwia obliczenie takich istotnych parametrów jak dobroć badanego rezonatora czy tłumienie jednostkowe przewodniczy.

W niniejszej pracy zajmujemy się analizą struktur zawierających jednorodnie prostopadłościennie wkładki dielektryczne. Do tej grupy struktur zaliczają się rozmaite modyfikacje linii typu „image” zwanej również falowodem zwierciadlanym, które konstruowane są w oparciu o falowód prostokątny zawierający jeden lub kilka dielektrycznych rdzeni. Do rozwiązania zagadnienia brzegowego zastosujemy opracowaną niedawno iteracyjną metodę rozkładu na funkcje własne (IMRW). Metoda ta, stosowana wcześniej z dużym powodzeniem do badania struktur światłowodowych [1], prowadzi do uzyskania pełnofalowego rozwiązania zagadnienia własnego przy równoczesnym uniknięciu problemów związanych z konstrukcją i rozwiązaniem układów równań liniowych, charakteryzujących dotychczas stosowane techniki.

## 2. OPIS ALGORYTMU IMRW

W swojej podstawowej postaci IMRW umożliwia rozwiązanie następującego problemu własnego:

$$(\mathbf{T} - \lambda)u = (\mathbf{L} - \mathbf{F})u - \lambda u = 0 \quad (1)$$

gdzie  $\mathbf{T}$  jest operatorem działającym w przestrzeni Hilberta, dla którego można dokonać dekompozycji  $\mathbf{T} = (\mathbf{L} - \mathbf{F})$  takiej, że  $\mathbf{L}$  jest operatorem samosprężonym a  $\mathbf{F}$  jest operatorem relatywnie zwartym względem operatora  $\mathbf{L}$ . W problemie (1)  $u$  i  $\lambda$  są odpowiednio szukaną funkcją i wartością własną operatora  $\mathbf{T}$ .

Przed przystąpieniem do rozwiązywania zagadnienia (1) konieczne jest wyznaczenie zbioru wartości własnych i funkcji własnych operatora  $\mathbf{L}$ . Jeżeli przyjmiemy, że  $i$ -ta wartość własna oraz  $i$ -ta funkcja własna operatora  $\mathbf{L}$  oznaczone są przez odpowiednio  $\Lambda_i$  i  $f_i$ , to algorytm wyznaczania rozwiązania problemu własnego (1) przedstawić można w postaci następującej pętli iteracyjnej:

Krok 0

A) Określenie zerowego przybliżenia  $u_0, \lambda_0$  takiego, że  $\mathbf{F}u_0 \neq 0$ , zaś  $\lambda_0$  jest wartością regularną operatora  $\mathbf{L}$ .

Krok  $k$ -ty

B) Znalezienie współczynników rozwinięcia  $k$ -tego przybliżenia funkcji  $u$  w szereg funkcji własnych operatora  $\mathbf{L}$ . Rozwinięcie to można zapisać jako:

$$\tilde{u}^{(k)} = \sum_i c_i^{(k)} f_i \quad (2)$$

gdzie

$$c_i^{(k)} = \frac{\langle \mathbf{F}u^{(k-1)}, f_i \rangle}{\Lambda_i - \lambda^{(k-1)}} \quad (3)$$

przy czym symbol  $\langle \cdot, \cdot \rangle$  oznacza iloczyn skalarny.

C) Normalizacja  $\tilde{u}^{(k)}$  polegająca na normalizacji współczynników rozwinięcia  $c_i^{(k)}$ .

$$u^{(k)} = \sum_i d_i^{(k)} f_i \quad (4)$$

gdzie:

$$d_i^{(k)} = \frac{c_i^{(k)}}{\sqrt{\sum_i |c_i^{(k)}|^2}}. \quad (5)$$

D) Wyznaczenie  $k$ -tego przybliżenia wartości własnej w oparciu o formułę

$$\lambda^{(k)} = \lambda^{(k-1)} + \sum_i (A_i - \lambda^{(k-1)}) |d_i^{(k)}|^2 - \sum_i d_i^{*(k)} \langle \mathbf{F} u^{(k)}, f_i \rangle \quad (6)$$

lub

$$\lambda^{(k)} = \sum_i A_i |d_i^{(k)}|^2 - \sum_i d_i^{*(k)} \langle \mathbf{F} u^{(k)}, f_i \rangle.$$

Kroki  $B, C, D$  powtarzane są iteracyjnie do czasu uzyskania stabilności rozwiązania, która może być zdefiniowana jako spełnienie poniższych kryteriów:

$$\left| \frac{\lambda^{(k)} - \lambda^{(k-1)}}{\lambda^{(k)}} \right| \leq \varepsilon_\lambda; \quad \left| \|\tilde{u}^{(k)}\| - 1 \right| \leq \varepsilon_u \quad (7 \text{ a—b})$$

### 3. ZASTOSOWANIE IMRW DO ANALIZY PROWADNIC MIKROFALOWYCH

Zajmiemy się obecnie zastosowaniem IMRW do badania przewodnic zbudowanych w oparciu o falowód prostokątny, będących modyfikacją linii „image”. Falowód prostokątny zawierający prostokątne izotropowe wkładki różniące się między sobą przenikalnością elektryczną oraz wymiarami można traktować jako jednorodną w kierunku propagacji przewodnicę cylindryczną wypełnioną materiałem, którego przenikalność elektryczna jest funkcją położenia. Propagację wzdłuż osi z harmonicznie zmiennej w czasie fali elektromagnetycznej opisuje w takiej strukturze następujące równanie falowe:

$$\nabla_t^2 \vec{H}_t + k_0^2 \varepsilon(x, y) \vec{H}_t + \frac{1}{\varepsilon(x, y)} [\nabla_t \varepsilon(x, y) \times (\nabla_t \times \vec{H}_t)] - \beta^2 \vec{H}_t = 0 \quad (8)$$

gdzie  $k_0, \beta$  są odpowiednio liczbą falową próżni oraz współczynnikiem propagacji zaś indeks  $t$  określa składowe poprzeczne.

Równanie falowe (8) można przedstawić w postaci (1) przy czym  $\lambda = \beta^2$ ,  $u = \vec{H}_t = [H_x, H_y]$  oraz

$$\mathbf{L} = \nabla_t^2 + k_0^2$$

$$\mathbf{F} = -(\varepsilon(x, y) - 1)k_0^2 - \frac{1}{\varepsilon(x, y)} [\nabla_t \varepsilon(x, y) \times (\nabla_t \times \cdot)]. \quad (9)$$

Możemy przyjąć za [2], iż dziedziną operatora  $\mathbf{T}$  jest zdefiniowana w przestrzeni Hilberta  $\mathcal{H}$  wektorowych funkcji zespolonych określonych na obszarze  $\Omega$  będącym przekrojem poprzecznym przewodnicy, całkowalnych z kwadratem, poprzez zażądanie spełnienia przez pole  $H$  na brzegu  $\partial\Omega$  warunków wymuszonych istnieniem idealnego ekranu elektrycznego. Formalnie dziedzinę operatora  $\mathbf{T}$  zapisać można w następujący sposób:

$$D(\mathbf{T}) := \{ \vec{H}_t \in W_0^1(\Omega) : \vec{H}_t \in C^2 \text{ do brzegu } \partial\Omega \\ \vec{n} \cdot \vec{H}_t = \vec{n} \cdot (\nabla_t \times \vec{H}_t) = 0|_{\partial\Omega} \} \otimes C^2 \quad (9a)$$

gdzie  $W_0^1(\Omega)$  jest lokalną przestrzenią Sobolewa zaś  $\vec{n}$  jest wektorem jednostkowym prostopadłym do brzegu  $\partial\Omega$ . Można pokazać [2], iż przy tej dziedzinie operator  $\mathbf{F}$  jest relatywnie zwarty i nieskończenie mały względem operatora  $\mathbf{L}$ .

Rozpisując w układzie współrzędnych prostokątnych operacje różniczkowe występujące w (9) otrzymujemy.

$$\mathbf{L} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k_0^2$$

$$\mathbf{F}_x = -[(\varepsilon(x, y) - 1)k_0^2]_x - \frac{1}{\varepsilon(x, y)} \frac{\partial}{\partial y} \varepsilon(x, y) \left[ -\frac{\partial}{\partial y} (\cdot); \frac{\partial}{\partial x} (\cdot) \right]$$

$$\mathbf{F}_y = -[(\varepsilon(x, y) - 1)k_0^2]_y - \frac{1}{\varepsilon(x, y)} \frac{\partial}{\partial x} \varepsilon(x, y) \left[ \frac{\partial}{\partial y} (\cdot); -\frac{\partial}{\partial x} (\cdot) \right]. \quad (10 \text{ a—c})$$

Uzyskany operator jest operatorem działającym na funkcje wektorowe. Oznaczmy przez  $h_x^i, h_y^i$  oraz  $\Lambda_x^i, \Lambda_y^i$  odpowiednio  $i$ -tą funkcję i wartość własną operatora  $\mathbf{L}$  działającego na składową pola  $H$  w kierunku  $x$  bądź  $y$ . Przy tych oznaczeniach zależności (2—6) przyjmują postać (indeksy przy współczynnikach  $c^i, d^i$  odnoszące się do numeru kroku pominięte dla większej czytelności):

$$H_x^{(k)} = \sum_i c_x^i h_x^i \quad (11)$$

$$H_y^{(k)} = \sum_i c_y^i h_y^i \quad (12)$$

$$c_x^i = \frac{\langle \mathbf{F}_x \vec{H}_t^{(k-1)}, h_x^i \rangle}{\Lambda_x^i - \beta^2} \quad (13)$$

$$c_y^i = \frac{\langle \mathbf{F}_y \vec{H}_t^{(k-1)}, h_y^i \rangle}{\Lambda_y^i - \beta^2} \quad (14)$$

oraz

$$d_x^i = \frac{c_x^i}{\sqrt{\sum_i (|c_x^i|^2 + |c_y^i|^2)}} \quad (15)$$

$$d_y^i = \frac{c_y^i}{\sqrt{\sum_i (|c_x^i|^2 + |c_y^i|^2)}} \quad (16)$$

$$\beta^2 = \sum_i (\Lambda_x^i |d_x^i|^2 + \Lambda_y^i |d_y^i|^2) - \sum_i (d_x^{i*} \langle \mathbf{F}_x \vec{H}_t^{(k)}, h_x^i \rangle + d_y^{i*} \langle \mathbf{F}_y \vec{H}_t^{(k)}, h_y^i \rangle). \quad (17)$$

Iloczyn skalarny występujący w powyższych wyrażeniach rozumiany jest w sensie iloczynu w przestrzeni  $L^2$  tzn:

$$\langle u, v \rangle = \int_{\Omega} uv^* d\Omega \quad (18)$$

gdzie  $\Omega$  jest obszarem przekroju poprzecznego przewodnicy.

### 3.1. ANALITYCZNE WYZNACZENIE ILOCZYNÓW SKALARNYCH

Ponieważ przewodnica wypełniona jest szeregiem dielektryków, jej przekrój poprzeczny można podzielić na szereg prostokątnych podobszarów, wewnątrz których przenikalność elektryczna jest stała. Jeżeli przyjmiemy, że w  $\xi$ -tym podobszarze obejmującym wnętrze prostokąta  $\Omega_\xi: x_\xi \leq x \leq x_{\xi+1}; y_\xi \leq y \leq y_{\xi+1}$ , znajduje się ośrodek, który można scharakteryzować przy pomocy składowej (w ogólności zespolonego) przenikalności elektrycznej  $\varepsilon_\xi$  to funkcja  $\varepsilon(x, y)$  może być zapisana w postaci:

$$\varepsilon(x, y) = 1 + \sum_{\xi}^N (\varepsilon_\xi - 1) \chi_\xi(x, y) \quad (19)$$

gdzie  $N$  jest liczbą podobszarów zaś  $\chi_\xi(x, y)$  jest funkcją charakterystyczną obszaru  $\Omega_\xi$  zdefiniowaną następująco:

$$\chi_\xi(x, y) = \chi_\xi(x) \chi_\xi(y) = \begin{cases} 1 & x, y \in \Omega_\xi \\ 0 & x, y \notin \Omega_\xi \end{cases} \quad (20)$$

Operator  $\mathbf{L}$  jest operatorem Helmholtza określonym na prostokącie ograniczonym ściankami elektrycznymi falowodu. Funkcje własne operatora  $\mathbf{L}$  są rozwiązaniami równania Helmholtza dla jednorodnego falowodu prostokątnego. Tworzą one szereg funkcji harmoniczych, zatem wyznaczenie iloczynów skalarnych występujących w

(13, 14) i (17) możliwe jest na drodze analitycznej. Jedyne problemy stanowią nieciągłości funkcji  $\varepsilon(x, y)$ .

Podstawiając (19) do (10b—c) i działając operatorem  $\mathbf{F}$  na pole  $\vec{H}_t$  otrzymamy następujące wyrażenia:

$$\begin{aligned} \mathbf{F}_x \vec{H}_t &= -(\varepsilon(x, y) - 1)k_0^2 H_x - \frac{1}{\varepsilon(x, y)} \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(x) \frac{\partial}{\partial y} \chi_{\xi}(y) \left[ -\frac{\partial}{\partial y} H_x + \frac{\partial}{\partial x} H_y \right] \\ \mathbf{F}_y \vec{H}_t &= -(\varepsilon(x, y) - 1)k_0^2 H_y - \frac{1}{\varepsilon(x, y)} \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(y) \frac{\partial}{\partial x} \chi_{\xi}(x) \left[ \frac{\partial}{\partial y} H_x - \frac{\partial}{\partial x} H_y \right]. \end{aligned} \quad (21 \text{ a—b})$$

Posługując się teorią dystrybucji można obliczyć pochodną uogólnioną funkcji  $\chi_{\xi}(x)$  i  $\chi_{\xi}(y)$ .

$$\frac{\partial}{\partial v} \chi_{\xi}(v) = \delta(v_{\xi}) - \delta(v_{\xi+1}) \quad (22)$$

gdzie  $\delta(v_{\xi})$  jest funkcją delta Diraca określoną w punkcie  $v_{\xi}$ .

Pojawienie się w operatorze  $\mathbf{F}$  funkcji  $\delta$  Diraca powoduje, iż wyznaczenie iloczynów skalarnych  $\langle \mathbf{F}u, f_i \rangle$  wymaga specjalnej uwagi. Problem sprowadza się do obliczenia całek następującej postaci:

$$A \int \frac{(\varepsilon_{\xi} - 1) \delta(v - v_{\xi})}{1 + (\varepsilon_{\xi} - 1) \mathbf{h}(v - v_{\xi})} g(v) dv, \quad (23)$$

gdzie  $\mathbf{h}(v - v_{\xi})$  jest funkcją Heaviside'a zdefiniowaną następująco:

$$\mathbf{h}(v) = \begin{cases} 0 & v < 0 \\ 1 & v > 0 \end{cases} \quad (24)$$

Przy obliczaniu (23) nie można stosować bezpośrednio twierdzeń dotyczących całkowania funkcji uogólnionych, ponieważ w punkcie występowania funkcji  $\delta$  funkcja podcałkowa zawiera nieciągłość w mianowniku. Usuniemy tę nieciągłość przyjmując, iż  $\mathbf{h}(0) = \frac{1}{2}$ . Stosując twierdzenie o całkowaniu dystrybucji otrzymujemy:

$$A \int \frac{(\varepsilon_{\xi} - 1) \delta(v - v_{\xi})}{1 + (\varepsilon_{\xi} - 1) \mathbf{h}(v - v_{\xi})} g(v) dv = \frac{A(\varepsilon_{\xi} - 1)g(v_{\xi})}{1 + 0,5(\varepsilon_{\xi} - 1)} = \frac{2A(\varepsilon_{\xi} - 1)}{(\varepsilon_{\xi} + 1)} g(v_{\xi}) \quad (25)$$

Posługując się formułą (25) oraz klasycznymi regułami całkowania można obliczyć analitycznie wszystkie składniki iloczynów skalarnych  $\langle \mathbf{F}u, f_i \rangle$  a w konsekwencji sprowadzić algorytm IMRW do prostego sumowania. Niezbędne zależności, umożliwiające analizę rozważanej klasy przewodnic mikrofalowych, podane są w dodatku A.

#### 4. UOGÓLNIENIE ALGORYTMU IMRW

W celu rozszerzenia klasy problemów, których rozwiązanie można znaleźć przy pomocy IMRW zajmijmy się obecnie uogólnionym zagadnieniem własnym postaci:

$$(\mathbf{L} - \mathbf{F})u - (\mathbf{G} + 1)\lambda u = 0 \quad (26)$$

gdzie  $\mathbf{G}$  jest funkcją gładką o nośniku zwartym. Dla tak zdefiniowanego problemu poszczególne kroki procesu IMRW będą przedstawiały się w następujący sposób:

Krok 0

A) Określenie zerowego przybliżenia  $u_0, \lambda_0$

Krok  $k$ -ty

B) Znalezienie współczynników rozwinięcia  $k$ -tego przybliżenia funkcji  $u$  w szereg funkcji własnych operatora  $\mathbf{L}$ .

$$c_i^{(k)} = \frac{\langle \mathbf{F}u^{(k-1)}, f_i \rangle + \lambda^{(k-1)} \langle \mathbf{G}u^{(k-1)}, f_i \rangle}{\Lambda_i - \lambda^{(k-1)}}. \quad (27)$$

C) Normalizacja współczynników rozwinięcia  $c_i^{(k)}$ .

$$d_i^{(k)} = \frac{c_i^{(k)}}{\sqrt{\sum_i |c_i^{(k)}|^2}}. \quad (28)$$

D) Wyznaczenie  $k$ -tego przybliżenia wartości własnej

$$\lambda^{(k)} = \frac{\sum_i \Lambda_i |d_i^{(k)}|^2 - \sum_i d_i^{*(k)} \langle \mathbf{F}u^{(k)}, f_i \rangle}{\sum_i d_i^{*(k)} \langle \mathbf{G}u^{(k-1)}, f_i \rangle + 1}. \quad (29)$$

#### 5. ZASTOSOWANIE IMRW DO ANALIZY REZONATORÓW DIELEKTRYCZNYCH

Zajmijmy się obecnie zastosowaniem IMRW do rozwiązania trójwymiarowego problemu brzegowego jakim jest wyznaczenie częstotliwości drgań własnych rezonatora dielektrycznego umieszczonego w prostopadłościennym wnętrzu. Przyjmijmy, że  $\varepsilon$  jest funkcją (separowalną) wszystkich trzech zmiennych przestrzennych. Przy tym założeniu otrzymamy z równań Maxwella, po wyrugowaniu pola  $\vec{E}$ , następujące równanie różniczkowe:

$$\nabla^2 \vec{H} + k_0^2 \varepsilon(x, y, z) \vec{H} + \frac{1}{\varepsilon(x, y, z)} [\nabla \varepsilon(x, y, z) \times (\nabla \times \vec{H})] = 0. \quad (30)$$

Równanie powyższe można przedstawić w postaci uogólnionego problemu własnego przyjmując  $\lambda = -k_0^2, u = \vec{H} = [H_x, H_y, H_z]$  oraz

$$\mathbf{L} = \nabla^2$$

$$\mathbf{F} = -\frac{1}{\varepsilon(x, y, z)} [\nabla \varepsilon(x, y, z) \times (\nabla \times \cdot)]$$

$$\mathbf{G} = \varepsilon(x, y, z) - 1. \quad (31)$$

Oznaczenie wartości własnych  $k_0^2$  prowadzi do określenia częstotliwości rezonansowych układu. Przyjmując, podobnie jak w przypadku problemu dwuwymiarowego (prowadnica), że funkcja  $\varepsilon(x, y, z)$  ma postać:

$$\varepsilon(x, y, z) = 1 + \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(x) \chi_{\xi}(y) \chi_{\xi}(z) \quad (32)$$

i działając operatorem  $\mathbf{F}$  na pole  $\vec{H}$  otrzymamy następujące wyrażenia:

$$\begin{aligned} \mathbf{F}_x \vec{H} &= -\frac{1}{\varepsilon(x, y, z)} \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(x) \left\{ \chi_{\xi}(z) \frac{\partial}{\partial y} \chi_{\xi}(y) \left[ -\frac{\partial}{\partial y} H_x + \frac{\partial}{\partial x} H_y \right] + \right. \\ &\quad \left. \chi_{\xi}(y) \frac{\partial}{\partial z} \chi_{\xi}(z) \left[ \frac{\partial}{\partial z} H_x - \frac{\partial}{\partial x} H_z \right] \right\}, \\ \mathbf{F}_y \vec{H} &= -\frac{1}{\varepsilon(x, y, z)} \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(y) \left\{ \chi_{\xi}(z) \frac{\partial}{\partial x} \chi_{\xi}(x) \left[ \frac{\partial}{\partial y} H_x - \frac{\partial}{\partial x} H_y \right] - \right. \\ &\quad \left. \chi_{\xi}(x) \frac{\partial}{\partial z} \chi_{\xi}(z) \left[ \frac{\partial}{\partial y} H_z - \frac{\partial}{\partial z} H_y \right] \right\}, \\ \mathbf{F}_z \vec{H} &= -\frac{1}{\varepsilon(x, y, z)} \sum_{\xi}^N (\varepsilon_{\xi} - 1) \chi_{\xi}(z) \left\{ \chi_{\xi}(y) \frac{\partial}{\partial x} \chi_{\xi}(x) \left[ \frac{\partial}{\partial z} H_x - \frac{\partial}{\partial x} H_z \right] + \right. \\ &\quad \left. \chi_{\xi}(x) \frac{\partial}{\partial y} \chi_{\xi}(y) \left[ \frac{\partial}{\partial y} H_z - \frac{\partial}{\partial z} H_y \right] \right\}. \end{aligned} \quad (33 \text{ a—c})$$

Poszczególne wyrażenia w procesie iteracyjnym obliczyć można na podstawie następujących związków:

$$\begin{aligned} H_v^{(k-1)} &= \sum_i c_v^i h_v^i \\ c_v^i &\sim \frac{\langle \mathbf{F}_v \vec{H}^{(k-1)}, h_v^i \rangle - k_0^2 \langle (\varepsilon - 1) H_v^{(k-1)}, h_v^i \rangle}{A_v^i + k_0^2} \\ d_v^i &\sim \frac{c_v^i}{\sqrt{\sum_i (|c_x^i|^2 + |c_y^i|^2 + |c_z^i|^2)}} \\ \text{gdzie } v &= x, y, z \end{aligned} \quad (34-36)$$

oraz

$$\begin{aligned}
 k_0^2 = & \left\{ \sum_i (d_x^{i*} < \mathbf{F}_x \vec{H}^{(k)}, h_x^i > + d_y^{i*} < \mathbf{F}_y \vec{H}^{(k)}, h_y^i > + d_z^{i*} < \mathbf{F}_z \vec{H}^{(k)}, h_z^i >) \right. \\
 & \left. - \sum_i (A_x^i |d_x^i|^2 + A_y^i |d_y^i|^2 + A_z^i |d_z^i|^2) \right\} / \\
 & \left\{ \sum_i (d_x^{i*} < (\varepsilon - 1) H_x^{(k)}, h_x^i > + d_y^{i*} < (\varepsilon - 1) H_y^{(k)}, h_y^i > \right. \\
 & \left. + d_z^{i*} < (\varepsilon - 1) H_z^{(k)}, h_z^i >) + 1 \right. \quad (37)
 \end{aligned}$$

Również w tym przypadku wszelkie niezbędne w procesie iteracyjnym wielkości można obliczyć w sposób analityczny. Z uwagi na nieciągłości funkcji  $\varepsilon(x, y, z)$  należy przy całkowaniu posłużyć się formułą (25). Końcowa postać wzorów (33—37) podana została w dodatku B.

## 6. WYNIKI

Posługując się analizą naszkicowaną powyżej przygotowaliśmy program numeryczny umożliwiający badanie charakterystyk dyspersyjnych oraz wyznaczanie rozkładu pola dla szerokiej klasy dielektrycznych przewodnic mikrofalowych. Wykonaliśmy szereg testów numerycznych porównując otrzymane wyniki z danymi opublikowanymi w literaturze [3]. Otrzymane rezultaty przedstawione są w tabeli.

Tabela 1

Współczynnik propagacji  $\beta(\text{mm}^{-1})$  (odwrócony fal. zwierciadlany [3],  $f = 79,4 \text{ GHz}$ )

| Rodzaj     | [3]    | IMRW (25 f-cji) | IMRW (81 f-cji) | IMRW (289 f-cji) |
|------------|--------|-----------------|-----------------|------------------|
| $E_{11}^p$ | 2.9873 | 2.95346         | 2.9925          | 2.99362          |
| $H_{11}^p$ | 2.7203 | 2.67341         | 2.72067         | 2.72115          |

Przedstawione w tabeli wyniki pozwalają stwierdzić, iż rozwiązanie uzyskiwane przy pomocy IMRW stabilizuje się, z dokładnością nie gorszą niż 0.2%, w pobliżu rezultatu otrzymanego w oparciu o metodę dopasowania rodzajów, która powszechnie uważana jest za metodę najbardziej dokładną. Potwierdza to poprawność zarówno analizy jak i programu komputerowego.

Kolejne dwie tabele przedstawiają porównanie wyników i czasów obliczeń dla IMRW i metody rodzajów sprzężonych [4], w której problem brzegowy zamieniony zostaje na macierzowe zagadnienie własne. Rezultaty dotyczą obliczeń wykonanych na komputerze klasy IBM AT 12 MHz 0Ws z koprocesorem INTEL 80287—10. W przypadku metody rodzajów sprzężonych wyniki zaczerpnięto z pracy [4] a podany

czas dotyczy jedynie samego obliczenia współczynników propagacji. Wyznaczanie wektorów własnych, niezbędnych dla określenia rozkładu pola, dodatkowo wydłużyłoby obliczenia w metodzie rodzajów sprzężonych.

Rezultaty i czasy podane w Tabeli 3 dotyczą rozwiązań uzyskanych po 8 lub 9 krokach iteracyjnych. Zbieżność zdefiniowana była jako spełnienie kryterium (7a) przy czym  $\varepsilon_\lambda = 10^{-4}$ . W momencie zatrzymania procesu iteracyjnego norma rozwiązania

Tabela 2

Metoda rodzajów sprzężonych [4] (linia typu „image”,  $k_0a = 25$ )

| Liczba funkcji ( $N$ )               | 32    | 50    | 72    | 98    |
|--------------------------------------|-------|-------|-------|-------|
| $\beta k_0^{-1}$ (rozd. $E_{11}^y$ ) | 1.527 | 1.529 | 1.529 | 1.529 |
| $\beta k_0^{-1}$ (rozd. $E_{21}^y$ ) | 1.452 | 1.456 | 1.458 | 1.458 |
| Czas obliczeń* [s]                   | 28    | 68    | 180   | 420   |

\* Wyznaczenie jedynie współczynników propagacji

Tabela 3

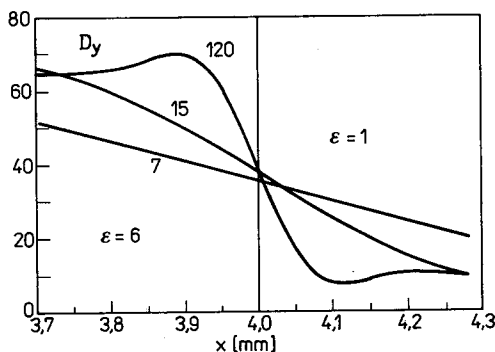
Metoda IMRW (linia typu „image”,  $k_0a = 25$ )

| Liczba funkcji ( $N$ )               | 32    | 50    | 72    | 98    | 200   |
|--------------------------------------|-------|-------|-------|-------|-------|
| $\beta k_0^{-1}$ (rozd. $E_{11}^y$ ) | 1.528 | 1.529 | 1.530 | 1.530 | 1.530 |
| Czas obliczeń** [s]                  | 6     | 8     | 11    | 14    | 34    |
| $\beta k_0^{-1}$ (rozd. $E_{21}^y$ ) | 1.452 | 1.453 | 1.455 | 1.455 | 1.456 |
| Czas obliczeń** [s]                  | 6     | 9     | 14    | 20    | 50    |

\*\* We wszystkich przypadkach zbieżność przy  $\varepsilon_\lambda = 10^{-4}$  (patrz 7a) po ośmiu lub dziewięciu iteracjach.

nie różniła się od jedności o więcej niż  $2 \cdot 10^{-4}$ . Program dynamicznie określał, które iloczyny skalarne są istotne dla określenia współczynników rozwinięcia i stąd czasy dla poszczególnych rodzajów różnią się między sobą. Na czas obliczeń może mieć również wpływ przybliżenie początkowe oraz ustalane arbitralnie parametry procesu iteracyjnego. Przedstawione wyniki dają podstawę twierdzić, iż w przypadku wyznaczania współczynników propagacji rodzajów podstawowych, metoda IMRW jest znacznie efektywniejsza od metody rodzajów sprzężonych. Warto podkreślić, iż główną zaletą IMRW nie tyle jest szybkość obliczeń co praktycznie brak ograniczeń związanych z obszarem pamięci operacyjnej komputera. O ile metoda IMRW wymaga pamiętania zaledwie  $4N$  współczynników, metody macierzowe posługują się tabicami zawierającymi co najmniej  $4N^2$  elementów.

Na koniec przeprowadzono test numeryczny mający na celu sprawdzenie prawidłowości obliczenia pola elektromagnetycznego. Rys. 1 ilustruje typową sytuację. Kolejne krzywe odpowiadają maksymalnemu numerowi harmonicznej przestrzennej w kierunku  $x$ .



Rys. 1 Zachowanie się składowej stycznej wektora indukcji elektrycznej na granicy między dwoma ośrodkami ( $\epsilon_1 = 6, \epsilon_2 = 1$ , granica w przekroju  $x = 4$  mm). Parametrem jest liczba harmonicznych przestrzennych w kierunku  $x$ .

Z rysunku można wywnioskować, iż uzyskanie poprawnego obrazu rozkładu pola wymaga uwzględnienia znacznej liczby funkcji własnych. Ponieważ liczba funkcji własnych jest iloczynem maksymalnej liczby harmonicznych w kierunku  $x$  i  $y$  zatem uzyskanie dokładnego obrazu pola jest praktycznie niemożliwe w tych metodach macierzowych, w których rozwiązanie budowane jest przy pomocy tych samych funkcji własnych co w IMRW.

## DODATEK A

W niniejszym dodatku podane zostaną związki niezbędne przy analizie przewodnic mikrofalowych zbudowanych w oparciu o falowód prostokątny zawierający rdzenie dielektryczne.

### 1) FUNKCJE I WARTOŚCI WŁASNE OPERATORA $L$ .

Obszar  $\Omega$  jest prostokątem ograniczonym przez ścianki elektryczne falowodu o wymiarach  $b$  na  $a$ . Ponieważ operator  $L$  jest operatorem Helmholtza to jego funkcje własne i wartości własne dane są przez następujące związki:

$$h_x^i = A_{mn} \sin \frac{m\pi x}{b} \cos \frac{n\pi y}{a} \quad n = 0, 1, \dots, m = 1, 2, \dots$$

$$h_y^i = B_{kl} \cos \frac{k\pi x}{b} \sin \frac{l\pi y}{a} \quad k = 0, 1, \dots, l = 1, 2, \dots$$

$$\begin{aligned}
 A_x^i &= k_0^2 - \left(\frac{m\pi}{b}\right)^2 - \left(\frac{n\pi}{a}\right)^2 \\
 A_y^i &= k_0^2 - \left(\frac{k\pi}{b}\right)^2 - \left(\frac{l\pi}{a}\right)^2
 \end{aligned} \quad (\text{A.1} - \text{A.4})$$

Stałe  $A_{mn}$ ,  $B_{kl}$  są współczynnikami normalizacyjnymi

$$\begin{aligned}
 A_{mn} &= \sqrt{\frac{2(2 - \delta_{0n})}{ab}} \\
 B_{kl} &= \sqrt{\frac{2(2 - \delta_{0k})}{ab}}
 \end{aligned} \quad (\text{A.5} - \text{A.6})$$

gdzie  $\delta_{ik}$  jest deltą Kroneckera.

Ponieważ w praktyce często spotykana jest symetria struktury względem płaszczyzny  $y = a/2$  opowiadająca umieszczeniu w tym miejscu ścianki elektrycznej lub magnetycznej, podamy również postać funkcji własnych dla obydwu tych sytuacji:

$$\begin{aligned}
 h_x^i &= A_{mn} \sin \frac{m\pi x}{b} \cos \frac{2n\pi y}{a} \quad n = 0, 1, \dots, m = 1, 2, \dots \\
 h_y^i &= B_{kl} \cos \frac{k\pi x}{b} \sin \frac{2l\pi y}{a} \quad k = 0, 1, \dots, l = 1, 2, \dots
 \end{aligned} \quad (\text{A.7} - \text{A.8})$$

dla ścianki elektrycznej w płaszczyźnie  $y = a/2$  oraz

$$\begin{aligned}
 h_x^i &= A_{mn} \sin \frac{m\pi x}{b} \cos \frac{(2n+1)\pi y}{a} \quad n = 0, 1, \dots, m = 1, 2, \dots \\
 h_y^i &= B_{kl} \cos \frac{k\pi x}{b} \sin \frac{(2l-1)\pi y}{a} \quad k = 0, 1, \dots, l = 1, 2, \dots
 \end{aligned} \quad (\text{A.9} - \text{A.10})$$

dla ścianki magnetycznej.

## 2) ILOCZYNY SKALARNE

Przed podaniem zależności umożliwiających obliczenie iloczynów skalarnych występujące w związkach (13—14) oraz (17), oznaczmy przez  $[m]$ ,  $[n]$  pojawiające się w (A.1—A.5) dwuskładnikowe multiindeksy odpowiadające  $j$ -tej funkcji własnej. Podobnie dla funkcji  $i$ -tej wprowadźmy oznaczenia  $[q]$ ,  $[r]$ . Składowe funkcji własnej można wtedy zapisać jako:

$$h_x^i = A_{[q]} \sin k_{q_x} x \cos k_{q_y} y \quad h_y^i = B_{[r]} \cos k_{r_x} x \sin k_{r_y} y. \quad (\text{A.11})$$

Iloczyny skalarne mają przy tych oznaczeniach następującą postać:

$$\langle \mathbf{F}_v \bar{H}_t, h_v^i \rangle = \sum_j \sum_{\xi}^N (I_1^v + I_2^v + I_3^v) \quad (A.12)$$

$$v = x, y,$$

gdzie

$$\begin{aligned} I_1^x &= -d_x^j \{ k_0^2 (\varepsilon_\xi - 1) A_{[m]} A_{[q]} I_x^{ss} [m_x, q_x] I_y^{cc} [m_y, q_y] \} \\ I_2^x &= d_y^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} B_{[n]} A_{[q]} k_{n_x} I_x^{ss} [n_x, q_x] \Delta_y [n_y, q_y] \right\} \\ I_3^x &= -d_x^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} A_{[m]} A_{[q]} k_{m_y} I_x^{ss} [m_x, q_x] \Delta_y [m_y, q_y] \right\} \\ I_1^y &= -d_y^j \{ k_0^2 (\varepsilon_\xi - 1) B_{[n]} B_{[r]} I_x^{cc} [n_x, r_x] I_y^{ss} [n_y, r_y] \} \\ I_2^y &= -d_y^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} B_{[n]} B_{[r]} k_{n_x} \Delta_x [n_x, r_x] I_y^{ss} [n_y, r_y] \right\} \\ I_3^y &= d_x^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} A_{[m]} B_{[r]} k_{m_y} \Delta_x [m_x, r_x] I_y^{ss} [m_y, r_y] \right\}. \end{aligned} \quad (A.13—A.18)$$

Przy obliczaniu  $I_2^v, I_3^v, v = x, y$  posłużono się formułą (25). Znaczenia symboli  $I_v^{cc}, I_v^{ss}, \Delta_v, v = x, y$  podane są w Dodatku C.

Na koniec należy zwrócić uwagę, że w zależnościach A.13.—A.18 przy współczynnikach  $d_v^i$  pominięte są indeksy odnoszące się do numeru kroku. W  $k$ -tym kroku procesu iteracyjnego iloczyny skalarne potrzebne do obliczenia współczynników  $c_v^i$  (wzory 13—14) oblicza się na podstawie A.12—A.18 podstawiając współczynniki  $d_v^i$  wyliczone w kroku  $k-1$ . Natomiast iloczyn skalarny występujący w (17) oblicza się podstawiając w A.13—A.18 współczynniki  $d_v^{i(k)}$ .

## DODATEK B

W niniejszym dodatku podane zostaną związki niezbędne przy analizie rezonatorów mikrofalowych zbudowanych w oparciu o prostopadłościenną wnękę zawierającą wkładki dielektryczne.

### 1) FUNKCJE I WARTOŚCI WŁASNE OPERATORA L.

Obszar  $\Omega$  jest w tym przypadku prostopadłościannym w przestrzeni  $(x, y, z)$  ograniczonym przez ścianki elektryczne wnęki o wymiarach  $b$  na  $a$  na  $c$ . W wyrażeniach określających funkcje i wartości własne posłużymy się, różnymi dla poszczególnych składowych, trójskładnikowymi multiindeksami  $[q], [r], [s]$ . Przy tych oznaczeniach funkcje i wartości własne operatora  $L$  określone są przez związki:

$$h_x^i = A_{[q]} \sin k_{q_x} x \cos k_{q_y} y \cos k_{q_z} z \quad q_x = 1, 2, \dots, q_y, q_z = 0, 1, 2, \dots$$

$$\begin{aligned}
h_y^i &= B_{[r]} \cos k_{r_x} x \sin k_{r_y} y \cos k_{r_z} z & r_y = 1, 2, \dots, r_x, r_z = 0, 1, 2, \dots \\
h_z^i &= C_{[s]} \cos k_{s_x} x \cos k_{s_y} y \sin k_{s_z} z & s_z = 1, 2, \dots, s_x, s_y = 0, 1, 2, \dots \\
A_x^i &= -k_{q_x}^2 - k_{q_y}^2 - k_{q_z}^2 \\
A_y^i &= -k_{r_x}^2 - k_{r_y}^2 - k_{r_z}^2 \\
A_z^i &= k_{s_x}^2 - k_{s_y}^2 - k_{s_z}^2
\end{aligned} \tag{B.1—B.6}$$

gdzie

$$k_{u_x} = \frac{u_x \pi}{b}, \quad k_{u_y} = \frac{u_y \pi}{a}, \quad k_{u_z} = \frac{u_z \pi}{c}, \quad u = q, r, s. \tag{B.7}$$

Natomiast  $A_{[q]}, B_{[r]}, C_{[s]}$  są współczynnikami normalizacyjnymi:

$$\begin{aligned}
A_{[q]} &= A_{ijk} = \sqrt{\frac{2(2 - \delta_{0j})(2 - \delta_{0k})}{abc}}, \\
B_{[r]} &= B_{ijk} = \sqrt{\frac{2(2 - \delta_{0i})(2 - \delta_{0k})}{abc}}, \\
C_{[s]} &= C_{ijk} = \sqrt{\frac{2(2 - \delta_{0i})(2 - \delta_{0j})}{abc}},
\end{aligned} \tag{B.8—B.10}$$

gdzie  $\delta_{ik}$  jest deltą Kroneckera a indeksy  $i, j, k$  należy tu rozumieć jako składowe odpowiednich multiindeksów.

## 2) ILOCZYNY SKALARNE

Obecnie podane zostaną związki umożliwiające obliczenie iloczynów skalarnych występujących w (35) i (37). Przyjmując, że  $z$   $j$ -tą i  $i$ -tą funkcją własną skojarzone są odpowiednio multiindeksy  $[m], [n], [p]$  oraz  $[q], [r], [s]$  otrzymujemy:

$$\begin{aligned}
\langle \mathbf{G} H_v, h_v^i \rangle &= \sum_j \sum_{\xi}^N I_1^v \\
\langle \mathbf{F}_v \vec{H}_t, h_v^i \rangle &= \sum_j \sum_{\xi}^N (I_2^v + I_3^v + I_4^v + I_5^v) \\
v &= x, y, z,
\end{aligned} \tag{B.11—B.12}$$

gdzie

$$\begin{aligned}
I_1^x &= d_x^j \{ (\varepsilon_{\xi} - 1) A_{[m]} A_{[q]} I_x^{ss} [m_x, q_x] I_y^{cc} [m_y, q_y] I_z^{cc} [m_z, q_z] \} \\
I_2^x &= d_y^j \left\{ \frac{2(\varepsilon_{\xi} - 1)}{(\varepsilon_{\xi} + 1)} B_{[n]} A_{[q]} k_{n_x} I_x^{ss} [n_x, q_x] \Delta_y [n_y, q_y] I_z^{cc} [n_z, q_z] \right\} \\
I_3^x &= -d_x^j \left\{ \frac{2(\varepsilon_{\xi} - 1)}{(\varepsilon_{\xi} + 1)} A_{[m]} A_{[q]} k_{m_y} I_x^{ss} [m_x, q_x] \Delta_y [m_y, q_y] I_z^{cc} [m_z, q_z] \right\}
\end{aligned}$$

$$I_4^x = -d_x^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} A_{[m]} A_{[q]} k_{m_z} I_x^{ss}[m_x, q_x] I_y^{cc}[m_y, q_y] \Delta_z[m_z, q_z] \right\}$$

$$I_5^x = d_z^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} C_{[p]} A_{[q]} k_{p_x} I_x^{ss}[p_x, q_x] I_y^{cc}[p_y, q_y] \Delta_z[p_z, q_z] \right\}, \quad (\text{B.13—B.17})$$

$$I_1^y = d_y^j \{ (\varepsilon_\xi - 1) B_{[n]} B_{[r]} I_x^{cc}[n_x, r_x] I_y^{ss}[n_y, r_y] I_z^{cc}[n_z, r_z] \}$$

$$I_2^y = -d_y^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} B_{[n]} B_{[r]} k_{n_x} \Delta_x[n_x, r_x] I_y^{ss}[n_y, r_y] I_z^{cc}[n_z, r_z] \right\}$$

$$I_3^y = d_x^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} A_{[m]} B_{[r]} k_{m_y} \Delta_x[m_x, r_x] I_y^{ss}[m_y, r_y] I_z^{cc}[m_z, r_z] \right\}$$

$$I_4^y = -d_y^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} B_{[n]} B_{[r]} k_{n_x} I_x^{cc}[n_x, r_x] I_y^{ss}[n_y, r_y] \Delta_z[n_z, r_z] \right\}$$

$$I_5^y = d_z^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} C_{[p]} B_{[r]} k_{p_y} I_x^{cc}[p_x, r_x] I_y^{ss}[p_y, r_y] \Delta_z[p_z, r_z] \right\} \quad (\text{B.18—B.22})$$

$$I_1^z = d_z^j \{ (\varepsilon_\xi - 1) C_{[p]} C_{[s]} I_x^{cc}[p_x, s_x] I_y^{cc}[p_y, s_y] I_z^{ss}[p_z, s_z] \}$$

$$I_2^z = d_x^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} A_{[m]} C_{[s]} k_{m_z} \Delta_x[m_x, s_x] I_y^{cc}[m_y, s_y] I_z^{ss}[m_z, s_z] \right\}$$

$$I_3^z = -d_z^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} C_{[p]} C_{[s]} k_{p_x} \Delta_x[p_x, s_x] I_y^{cc}[p_y, s_y] I_z^{ss}[p_z, s_z] \right\}$$

$$I_4^z = -d_z^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} C_{[p]} C_{[s]} k_{p_y} I_x^{cc}[n_x, s_x] \Delta_y[p_y, s_y] I_z^{ss}[p_z, s_z] \right\}$$

$$I_5^z = d_y^j \left\{ \frac{2(\varepsilon_\xi - 1)}{(\varepsilon_\xi + 1)} B_{[n]} C_{[s]} k_{n_z} I_x^{cc}[n_x, s_x] \Delta_y[n_y, s_y] I_z^{ss}[n_z, s_z] \right\} \quad (\text{B.23—B.28})$$

Znaczenia symboli  $I_v^{cc}$ ,  $I_v^{ss}$ ,  $\Delta_v$ ,  $v = x, y, z$  podane są w Dodatku C.

## DODATEK C

W niniejszym dodatku podane są znaczenia symboli  $I_v^{cc}$ ,  $I_v^{ss}$ ,  $\Delta_v$ ,  $v = x, y, z$  pojawiających się w dodatkach A i B.

$$I_v^{ss}[o_v, u_v] = \int_{v_\xi}^{v_\xi+1} \sin k_{o_v} \sin k_{u_v} v dv$$

$$I_v^c[o_v, u_v] = \int_{v_\xi}^{v_\xi+1} \cos k_{o_v} v \cos k_{u_v} v dv$$

$$\Delta_v[o_v, u_v] = \sin k_{o_v} v_\xi \cos k_{u_v} v_\xi - \sin k_{o_v} v_{\xi+1} \cos k_{u_v} v_{\xi+1} \quad (C.1-C.3)$$

$v = x, y, z \quad o = m, n, p \quad u = q, r, s$

Praca finansowana przez Polską Akademię Nauk w ramach CPBP 02.02.3.2.

### BIBLIOGRAFIA

- 1 T.F. Jabłoński, M.J. Sowiński: *Analysis of dielectric guiding structures by the iterative eigenfunction expansion method. IEEE Trans. Microwave Theory Tech.*, vol. MTT-37, pp. 63—70, Jan. 1989
- 2 T.F. Jabłoński: *Iteracyjna metoda rozkładu na funkcje własne dla światłowodów cylindrycznych*, prace IPPT PAN, nr. 3, 1986
- 3 R. Mittra Y. Hou and V. Jamnejad: *Analysis of open dielectric waveguides using mode-matching technique and variational methods. IEEE Trans. Microwave Theory Tech.*, vol. MTT-28, pp. 36—43, Jan. 1980
- 4 K. Ogusu: *Numerical analysis of rectangular dielectric waveguide and its modifications. IEEE Trans. Microwave Theory Tech.*, vol. MTT-25, pp. 874—885, Nov. 1977
- 5 K. Bierwirth, N. Shultz, and F. Arndt: *Finite-difference analysis of rectangular dielectric waveguide structures. IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 1104—1113, Nov. 1986
- 6 W. McLevige, T. Itoh and R. Mittra: *New waveguide structures for millimeter and optical integrated circuits. IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 788—794, Oct. 1975
- 7 C. C. Su: *A combined method for dielectric waveguides using the finite-element technique and the surface integral equations method. IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 1140—1146, Nov. 1986

M. MROZOWSKI, J. MAZUR, T. JABŁOŃSKI

Application of the iterative eigenfunction expansion method to the analysis  
of microwave dielectric structures

### Summary

An analysis of microwave resonators and guides containing homogeneous dielectric slabs is presented. The boundary value problem is solved by means of the iterative eigenfunction expansion method (IEEM). Using the properties of generalized functions, all the expressions essential for the implementation of IEEM are derived, reducing the algorithm to a simple summation process. Numerical results obtained by IEEM are compared with published data, showing a very good agreement and a higher efficiency as compared with other rigorous approaches.

# Współczesne metody pomiaru anten mikrofalowych

WOJCIECH J. KRZYSZTOFIK

*Instytut Telekomunikacji i Akustyki, Politechnika Wrocławska*

*Otrzymano 1989.02.28*

*Autoryzowano do druku 1989.04.03*

Szybki rozwój systemów łączności, komunikacji morskiej i lądowej (zwłaszcza satelitarnej) wymaga dokładnego określenia podstawowych parametrów stosowanych w nich anten. Coraz bardziej krytyczne zaczynają być wymagania kompatybilnej współpracy różnych systemów radiokomunikacyjnych stosowanych przez różne służby, co szczególnie daje się zauważyć na ograniczonej powierzchni statków. W pierwszej części pracy zdefiniowano warunki jakie musi spełniać stanowisko pomiarowe, oraz zaprezentowano podstawowe typy stanowisk (w polu bliskim, pośrednim i dalekim). W drugiej części zdefiniowano tzw. polowe charakterystyki anten (zespoloną charakterystykę promieniowania, polaryzację, kierunkowość, zysk energetyczny) oraz przedstawiono schematy blokowe układów pomiarowych.

## WSTĘP

Antena jest urządzeniem przetwarzającym energię fali elektromagnetycznej prowadzonej w linii transmisyjnej na energię fali promieniowanej w swobodnej przestrzeni (antena nadawcza) i odwrotnie (antena odbiorcza). Ze względu na miejsce i rolę w systemie radiokomunikacyjnym, antena musi spełniać funkcję układu dopasowującego parametry urządzeń nadawczych i/lub odbiorczych do parametrów środowiska propagacyjnego.

Parametry elektryczne anteny zainstalowanej w systemie pracującym w określonych warunkach (na statkach powietrznych i morskich, w otoczeniu innych anten, na dachach budynków, itp.) różnią się, na ogół, znacznie od parametrów tej samej anteny umieszczonej w swobodnej przestrzeni. Jest to spowodowane udziałem otoczenia anteny w formowaniu się wypadkowego pola promieniowanego. Teoretyczne określenie tych zmian jest możliwe tylko przy daleko idących założeniach upraszczających i dostarcza tylko wyników jakościowych. Ostateczna weryfikacja parametrów anten musi być dokonana doświadczalnie przez pomiar bezpośredni lub na modelach.

Szybki rozwój systemów łączności i radionawigacji w komunikacji morskiej i lotniczej, a zwłaszcza satelitarnej, w których prawidłowy dobór zestawu, typów, rozmiarów i wzajemnej konfiguracji anten (ze względu na optymalne wykorzystanie ograniczonej powierzchni statku) wymaga dokładnego określenia parametrów anten, spowodował opracowanie nowych metod pomiarowych [1, 2, 5]. Oprócz konwencjo-

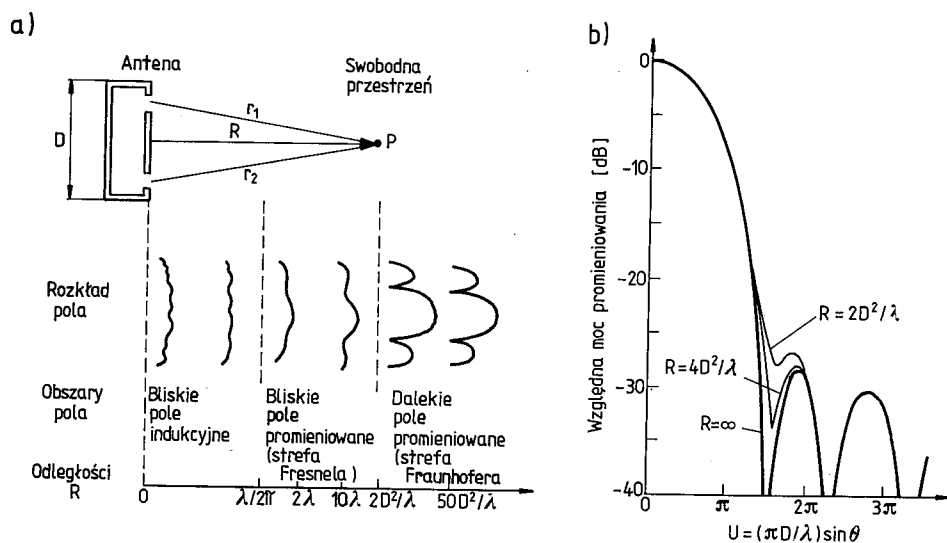
nalnych pomiarów w strefie dalekiej wykonuje się pomiary w strefie bliskiej i pośredniej (tzw. stanowiska zredukowane i rozogniskowane). Jako źródła pola elektromagnetycznego wykorzystuje się sygnały z radiogwiazd (metody radioastronomiczne) i ze sztucznych satelitów Ziemi (metody satelitarne). Do obserwacji pola promieniowanego, zwłaszcza w pobliżu anteny, stosuje się detektory ciekłokrystaliczne. Pomiary mniejszych anten i modeli struktur promieniujących wykonuje się w specjalnych komorach bezchowych. W celu maksymalnego skrócenia czasu oraz poprawienia dokładności pomiarów opracowano automatyczne, sterowane mikroprocesorami, stanowiska kontrolno-pomiarowe.

## Część I

### 1. WARUNKI POMIAROWE

#### 1.1. GEOMETRIA STANOWISKA POMIAROWEGO

Przestrzeń z polem elektromagnetycznym wytwarzanym przez antenę jest zwykle dzielona na trzy podobszary: (I) obszar pola indukcyjnego, (II) obszar bliskiego pola promieniowania, (III) obszar dalekiego pola promieniowania (rys. 1).



Rys. 1. Obszary pola elektromagnetycznego w otoczeniu anteny (a), wpływ odległości pomiarowej na kształt pomierzonej charakterystyki promieniowania anteny (b)

Pole elektromagnetyczne w poszczególnych obszarach ma różną strukturę i właściwości. W miarę oddalania się od anteny fala zmienia się z kulistej, o skomplikowanych zależnościach fazowych pomiędzy poszczególnymi składowymi pola, w płaską poprzeczną (TEM) — w obszarze dalekim.

Polowe charakterystyki anten (charakterystyka promieniowania, polaryzacja, kierunkowość, zys energetyczny) są definiowane dla dalekiego pola promieniowania (fala TEM). Badana antena powinna się więc znajdować w polu fali płaskiej źródła pola elektromagnetycznego. Przy pomiarze charakterystyk badane anteny (jeśli nie zawierają elementów nieodwracalnych) pracują najczęściej jako anteny odbiorcze, bez względu na ich rzeczywiste przeznaczenie.

W praktyce, w trakcie pomiarów, trudno jest spełnić kryterium pola dalekiego, zapewniające płaski kształt czoła fali oświetlającej aperturę badanej anteny. Dla założonego dopuszczalnego odchylenia fazy  $\Delta\varphi$  w aperturze  $D$  badanej anteny, minimalną odległość pomiarową  $R$  (promień strefy dalekiej) określamy ze wzoru [1]

$$R = \frac{\pi D^2}{4\Delta\varphi\lambda}. \quad (1.1)$$

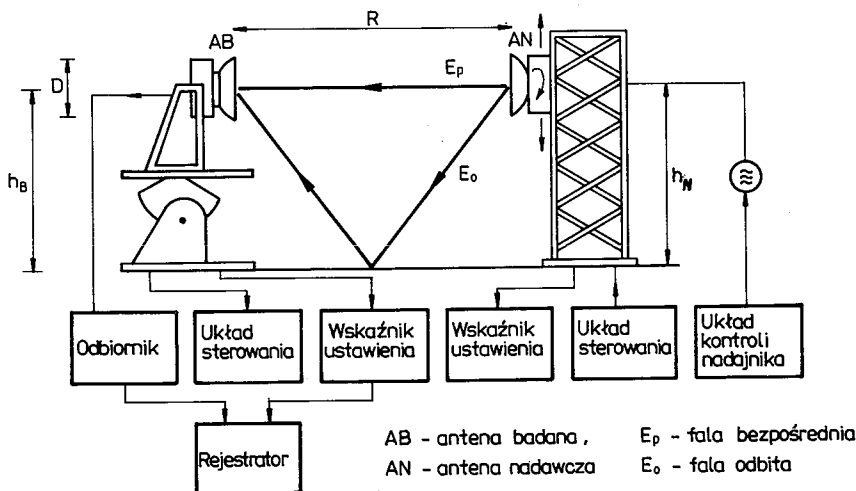
Przy pomiarach charakterystyk amplitudowych dopuszcza się odchylenie fazy  $\Delta\varphi = \pi/8$ , wówczas

$$R = 2D^2/\lambda. \quad (1.2)$$

Zależność ta jest słuszna dla anten dużych elektrycznie ( $D > \lambda$ , tzn. o dużych kierunkowościach), dla innych przyjmuje się  $R \geq 10 \lambda$ . W przypadku pomiarów charakterystyk fazowych odchylenia fazowe w aperturze badanej anteny mają duży wpływ na dokładność otrzymanych wyników. Przyjmując np.  $\Delta\varphi = \pi/90$  musimy zwiększyć odległość pomiarową ponad dziesięć razy, w stosunku do obliczonej ze wzoru (2).

Pomiary charakterystyk anten metodami klasycznymi są wykonywane w strefie dalekiej. Schemat i konfigurację stanowiska pomiarowego przedstawiono na rys. 2.

W ogólnym przypadku do badanej anteny docierają dwie fale: bezpośrednia ( $E_p$ ) i odbita ( $E_o$ ). Zjawisko to jest przyczyną powstawania interferencji pola, które w sposób



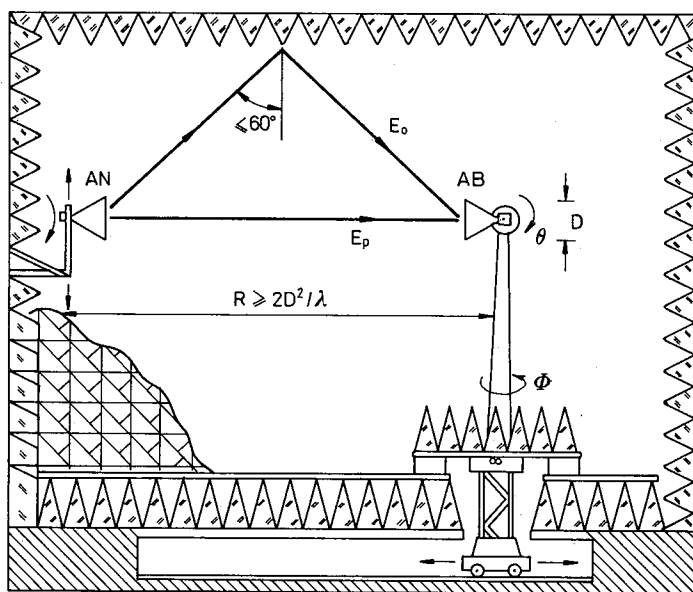
Rys. 2. Schemat blokowy stanowiska pomiarowego charakterystyk anten mikrofalowych

znaczący mogą zmieniać charakterystyki anteny (np. pole interferencyjne o poziomie  $-30$  dB może zmienić pomierzony zysk o  $\pm 0.25$  dB). Najczęściej stosowanymi metodami ograniczenia odbić w obszarze pomiarowym są: umieszczenie anten na wieżach lub wzniesieniach, zastosowanie kierunkowych anten nadawczych, dobór konfiguracji stanowiska pomiarowego, zapewniający wyeliminowanie powierzchni odbijających falę w kierunku badanej anteny, ustawienie przegród lub powierzchni pochłaniających fale odbite [5]. Czasami jednak, w celu zwiększenie poziomu sygnału dochodzącego do badanej anteny, tak dobiera się wysokości zawieszenia anten

$$h_N = \lambda R / (4h_B), \quad (1.3)$$

aby uzyskać sumowanie obu składowych ( $E_p$  i  $E_o$ ) pola (nie stosuje się tego przy pomiarach polaryzacji — fala odbita zmienia polaryzację).

Skuteczną metodą eliminacji wielodrogowości (interferencji) sygnałów docierających do badanej anteny, ograniczenia wpływów zakłóceń zewnętrznymi polami elektromagnetycznymi oraz ochrony przed wpływami atmosferycznymi jest umieszczenie stanowiska pomiarowego w komorze bezodbiciowej (bezechowej). Przykład komory bezechowej przedstawiono na rys. 3.



Rys. 3. Stanowisko pomiarowe w komorze bezechowej

Komory bezechowe są wykładane materiałem dielektrycznym i/lub magnetycznym (poliuretan pokryty grafitem, ferryt, itp.), w celu zamiany energii pola padającego na ściany w ciepło. W celu optymalnego tłumienia, pochłaniania i rozpraszania promieni fal dochodzących do ścian, wykładzina komory jest wykonywana w kształcie ostrosłupów.

Obecnie produkowane komory bezechowe umożliwiają wykonanie pomiarów anten w zakresie częstotliwości 0.1 do 100 GHz, z maksymalnym współczynnikiem

odbicia (fali prostopadłej do ściany) — 40 dB. Proporcje (długość — szerokość, najczęściej 2:1) komory, dobiera się ze względu na maksymalizację rozmiarów strefy pomiarowej, wolnej od odbić, z prawie płaskim kształtem czoła fali. W celu częstotliwościowego rozszerzenia zakresu pomiarowego stosuje się komory piramidalne [1, 5].

## 1.2. APARATURA POMIAROWA

Na rys. 2 przedstawiono przykładowe stanowisko do pomiaru charakterystyk anten. Podstawowymi blokami systemu pomiarowego są: anteny i urządzenia nadawcze, urządzenia odbiorcze, układy ustawiania i sterowania położenia anten, systemy rejestracji i przetwarzania danych pomiarowych. Źródło mocy mikrofalowej (nadajnik) powinno być wyposażone w układy kontroli poziomu sygnału wyjściowego, stabilizacji oraz wobulowania i/lub modulacji częstotliwości. W zakresie mikrofal (powyżej 1 GHz) jako nadawcze anteny pomiarowe stosuje się: dipole, anteny logarytmiczno-periodyczne, tuby oraz reflektory paraboliczne, oświetlane wcześniej wymienionymi antenami.

Odbiornikiem sygnału pomiarowego może być prosty detektor krystaliczny lub bolometr, załączony na zaciski badanej anteny. Przy pomiarach charakterystyk fazowych i polaryzacyjnych stosuje się odbiorniki heterodynowe lub homodynowe (woltomierze wektorowe, analizatory układów). Badana antena jest umieszczona na stole obrotowym (azymut — elewacja), sterowanym zdalnie przez urządzenie kontrolne i sprzężonym z urządzeniami rejestrującymi i/lub drukującymi (przetwarzającymi) wyniki pomiarów. Wykresy charakterystyk anten przedstawia się w prostokątnym lub biegunowym układzie współrzędnych za pomocą drukarki i/lub zapisuje na rejestratorze. Stanowisko pomiarowe powinno zapewnić dynamikę pomiarów (poziom zakłóceń, czułość przyrządów, itp.)  $0 \div 60$  dB ( $\geq 40$  dB).

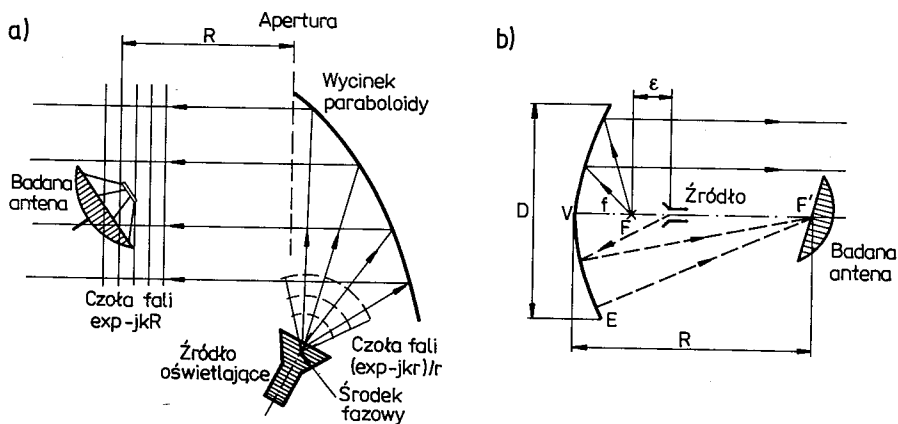
## 1.3. STANOWISKO POMIAROWE O ZREDUKOWANYCH ROZMIARACH

Rozmiary stanowiska pomiarowego, wynikające z warunku strefy dalekiej (1) lub (2) są często duże (np. dla  $D = 1$  m, przy  $f = 30$  GHz,  $R = 200$  m), co utrudnia wybór obszaru wolnego od odbić i przeszkód pomiędzy antenami lub wymaga budowy dużych komór (podraża koszty stanowiska). Prawie płaską falę elektromagnetyczną można wytworzyć w mniejszych odległościach od określanych z (2), stosując reflektory paraboliczne (stanowiska zwarte) lub wprowadzając rozogniskowanie anteny reflektorowej, wytwarzającej pole pomiarowe dla anteny badanej (stanowiska rozogniskowane).

Istotę budowy stanowiska zwanego można przeanalizować metodą optyki geometrycznej (rys. 4a). Zakłada się, że pole elektromagnetyczne jest wytwarzane przez reflektor paraboliczny, oświetlany, umieszczonym w jego ognisku, źródłem punktowym (liniowym). Kuliste (walcowe) czoło fali promieniowanej przez źródło, po odbiciu od reflektora jest formowane w czoło fali płaskiej. Promienie rozchodzą się od źródła przez reflektor w kierunku badanej anteny. Pole wytwarzane przez źródło nie zależy od odległości pomiędzy aperturą reflektora i badaną anteną. W przeciwieństwie

do pomiarów w strefie dalekiej, wzrost częstotliwości pomiarowej może poprawić parametry (równomierność rozkładu amplitud) obszaru pomiarowego bez zmiany wymiarów stanowiska.

W praktyce, występują pewne odstępstwa od optyki geometrycznej: nierównoległy przebieg promieni po odbiciu od reflektora (przy większych  $\lambda$ , niedokładnym wykonaniu powierzchni paraboloidy lub niewłaściwym wyznaczeniu położenia ogniska), depolaryzacja fali odbitej, sygnały niepożądane, spowodowane np. promieniowaniem źródła przez listki boczne, itp.



Rys. 4. Stanowiska pomiarowe o zredukowanych rozmiarach: zwarte (a) i rozogniskowane (b)

Projektując stanowisko zwarte należy wybrać reflektor o co najmniej trzykrotnie większej aperturze od badanej anteny, o dokładnie wykonanej powierzchni (około  $0.1 \lambda$  na każdy  $\lambda^2$  element powierzchni) i długiej ogniskowej (minimalizacja efektów depolaryzacji fali odbitej). Typowo, w zakresie częstotliwości (2 do 40) GHz można uzyskać cylindryczny obszar pomiarowy o wymiarach  $(1 \times 1,5)$  m [3]. Innym przykładem stanowiska pomiarowego o zmniejszonych rozmiarach jest obszar z reflektorem paraboloidalnym o przesuniętym ognisku, tzw. stanowisko rozogniskowane (rys. 4b).

Reflektor paraboliczny ma tę właściwość, że promienie wychodzące z ogniska i docierające do płaszczyzny apertury mają jednakową długość, a w dużych odległościach od niej tworzą falę płaską. W celu uzyskania takich warunków w punkcie F' odległym o R od wierzchołka V paraboli, należy przesunąć o  $\epsilon$  ognisko F paraboli (rys. 4b).

Położenie nowego ogniska określa się z przybliżonej zależności [4]

$$\epsilon \cong (0.05 \div 0.1)(f^2 + (D/4)^2)R \quad (1.4)$$

(f jest ogniskową reflektora).

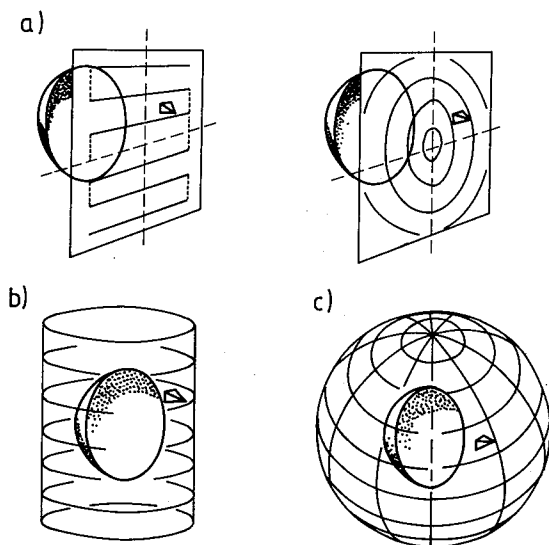
Odległości pomiarowe na stanowisku rozogniskowanym, w zależności od rozmiarów badanej anteny, mieszczą się w przedziale  $\left(\frac{D^2}{4\lambda}\right)$  do  $\left(\frac{2D^2}{\lambda}\right)$ . Na obu stanowiskach o

zmniejszonych rozmiarach można wykonywać pomiary anten o dużych kierunkowościach, ponieważ zasada skrócenia odległości strefy dalekiej jest słuszna tylko w niewielkim przedziale kątów wokół osi reflektora.

#### 1.4. POMIARY W POLU BLISKIM ANTENY

Dotychczas omówione metody pomiarowe polegały na określaniu rozkładu pola elektromagnetycznego w strefie dalekiej lub mniejszej (z zachowaniem warunków jak dla strefy dalekiej), na podstawie którego można określić polowe charakterystyki anten. Metody te, ze względu na wymagane odległości pomiarowe oraz konieczność umieszczenia badanej anteny na stole obrotowym, nie mogą być stosowane do pomiarów (na stanowiskach pomiarowych) dużych anten.

Alternatywnym rozwiązaniem, coraz szerzej stosowanym, jest pomiar w strefie bliskiego pola promieniowania badanej anteny. Polega on na wyznaczaniu rozkładu wektora (moduł i faza) pola elektrycznego na wybranej aperturze pomiarowej otaczającej antenę (rys. 5).



Rys. 5. Przykłady kształtów apertur pomiarowych, stosowanych przy pomiarze charakterystyk anten w polu bliskim: płaska (a), cylindryczna (b), sferyczna (c)

W zależności od wymiarów i kierunkowości badanej anteny wybiera się apertury pomiarowe w kształcie: płaszczyzny, walca lub sfery. Rozkład pola na aperturze mierzy się dla dwóch polaryzacji (pionowej —  $V$  i poziomej —  $H$ ) nadawczej sondy pomiarowej. Dla dowolnego punktu  $x, y$  na powierzchni apertury, oddalonej od badanej anteny o  $z = d$ , pomierzone napięcia na zaciskach sondy oznaczmy  $u_V$  i  $u_H$ , odpowiednio dla polaryzacji pionowej i poziomej. Pole w strefie dalekiej, promieniowane przez dowolną aperturę jest transformatą Fouriera rozkładu pola w aperturze i odwrotnie (teoria dyfrakcji). W przypadku płaskiej apertury pomiarowej (o rozmia-

rach X na Y), pole w strefie dalekiej (składowe  $E_\theta$  i  $E_\phi$ ) badanej anteny jest określone następującą zależnością [6]:

$$E_{\theta, \phi}(\Theta, \Phi) = A [I_H(\Theta, \Phi) E_{\theta, \phi}^V(\pi - \Theta, \Phi) - I_V(\Theta, \Phi) E_{\theta, \phi}^H(\pi - \Theta, \Phi)], \quad (1.5)$$

w której:

$E_{\theta, \phi}^V(\cdot)$  i  $E_{\theta, \phi}^H(\cdot)$  są charakterystykami promieniowania sondy (dla obu polaryzacji)

$$I_{V, H}(\Theta, \Phi) = \exp(jk_z d) \int_{-X/2}^{X/2} \int_{-Y/2}^{Y/2} u_{V, H}(d, x, y) \exp(j(k_x x + k_y y)) dx dy.$$

$$A = \sin \Theta \cos \Phi / \Delta(\Theta, \Phi),$$

$$\Delta(\Theta, \Phi) = E_\theta^H(\pi - \Theta, \Phi) E_\theta^V(\pi - \Theta, \Phi) - E_\phi^V(\pi - \Theta, \Phi) E_\phi^H(\pi - \Theta, \Phi),$$

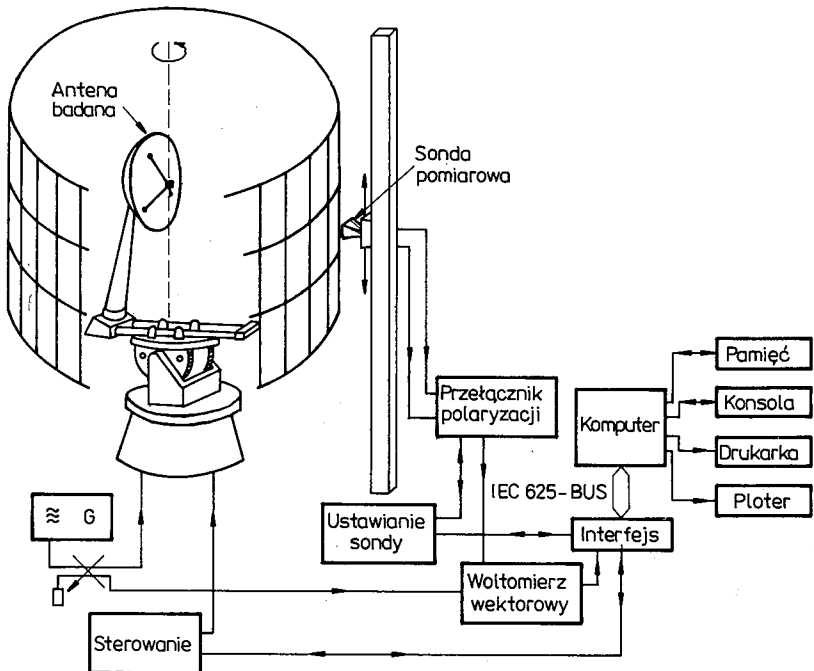
$$k_x = (2\pi \cos \Theta) / \lambda; \quad k_y = (2\pi \sin \Theta \sin \Phi) / \lambda; \quad k_z = [(2\pi / \lambda)^2 - k_x^2 - k_y^2]^{1/2},$$

Rozmiary apertury pomiarowej określają maksymalny kąt  $\Theta_m$ , w jakim mogą być wyliczone składowe pola dalekiego

$$\Theta_m \cong \text{ctg}(D - X) / 2d \quad (1.6)$$

gdzie: D i X — rozmiar badanej anteny i szerokość apertury.

Odległości między punktami sondowania na aperturze pomiarowej powinny być mniejsze od  $\lambda/2$  (np.  $\Delta x = \Delta y = \lambda/3$ , dla  $z = d = (2 \div 3)\lambda$ ). Ustawienie sondy powinno



Rys. 6. Schemat blokowy układu do pomiaru charakterystyk anten w bliskim polu promieniowanym z aperturą cylindryczną

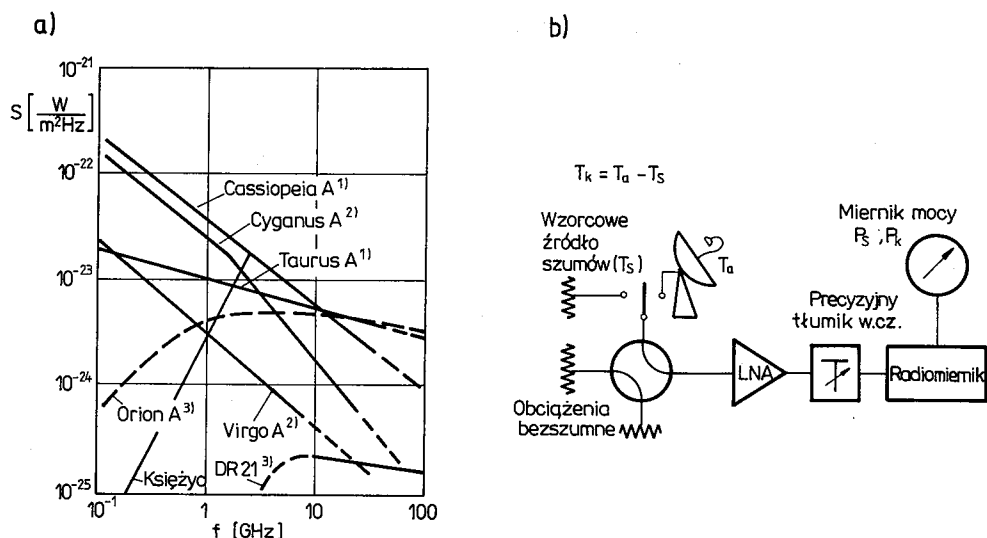
być określane z dokładnością  $\lambda/100$ . Dane pomiarowe ( $u_v$ ,  $u_H$ ) oraz współrzędne położenia ( $\Theta$ ,  $\Phi$ ,  $x$ ,  $y$ ,  $z$ ) sondy na aperturze są zapamiętywane i przetwarzane przez komputer. Całki dwuwymiarowe są obliczane za pomocą szybkiej transformaty Fouriera (FFT). Przykład stanowiska pomiarowego w strefie bliskiej przedstawiono na rys. 6. Wszystkie procedury sterowania obiektami na stanowisku i przetwarzania danych pomiarowych są sterowane przez komputer. Pomiary w strefie bliskiej są stosowane w zakresie częstotliwości ( $1 \div 65$ ) GHz. Ograniczenie w dolnym zakresie częstotliwości wynika z małych kierunkowości anten (duże apertury pomiarowe), natomiast w górnym z dokładności ustawienia sondy pomiarowej. Na podstawie wyników pomiarów w strefie bliskiej można określić wszystkie charakterystyki anteny z dokładnością porównywalną z pomiarami w strefie dalekiej lub lepszą. Ze względu na zmniejszenie rozmiarów stanowiska, pomiary mogą być wykonywane w komorach bezechowych. Niewątpliwie, do wad tej metody należy zaliczyć: znaczne skomplikowanie układu pomiarowego (wzorcowanie sond pomiarowych, precyzyjne pomiary ich położenia, komputer, oprogramowanie) oraz niemożność otrzymania charakterystyk badanej anteny w czasie rzeczywistym.

#### 1.5. POMIARY ANTEN W MIEJSCU ZAINSTALOWANIA

Wykonywanie pomiarów charakterystyk anten dużych (elektrycznie i fizycznie), na stanowiskach pomiarowych jest niepraktyczne a często — niemożliwe (duże rozmiary stanowiska, anteny niemobilne, anteny zainstalowane w określonym otoczeniu, itp.). Anteny naziemnych stacji satelitarnych, dużych radio-teleskopów, radio-nadawcze, itp. są, najczęściej, mierzone w miejscu ich zainstalowania [2, 5]. Jako źródła pola elektromagnetycznego, w tych przypadkach, wykorzystuje się naturalne promieniowanie słońca, księżyca, oraz radio-gwiazd (metody radio-astronomiczne), nadajniki umieszczone na statkach powietrznych, takich jak: balony, samoloty, helikoptery (np. metody śmigłowcowe) oraz sztucznych satelitach Ziemi (metody satelitarne).

##### Radio-astronomiczne metody pomiarowe

Niektóre planety emitują fale elektromagnetyczne, charakteryzujące się ściśle określoną widmową gęstością mocy, promieniowanej w niewielkich kątach brytowych. Położenie tych źródeł jest stałe i ściśle określone. Gęstości promieniowania kilku radio-gwiazd przedstawiono na rys. 7a [1]. Najczęściej wykorzystywanymi radio-gwiazdami są *Cassiopeia A*, *Cygnus A* i *Taurus A*. Dla anten, których szerokość wiązki głównej charakterystyki promieniowania mocy na poziomie  $-3$  dB jest większa od  $10'$ , radio-gwiazdy mogą być traktowane jako źródła punktowe. Zakłada się, że promieniowany przez nie sygnał ma polaryzację przypadkową, w związku z tym sprawność polaryzacji wynosi 0.5. Moc odebrana przez badaną antenę jest równa  $k \cdot T_a$  ( $k = 1.38 \cdot 10^{-23}$  J/K — stała Boltzmanna,  $T_a$  — zastępcza temperatura szumowa radio-gwiazdy, odniesienia do zacisków anteny).



Rys. 7. Widmowe gęstości promieniowania mocy kilku radio-gwiazd (quazarów) (a) 1 — SR pozostałość supernowa, 2 — RG radiowa droga mleczna, 3 — EM mgławica emisyjna  
b) schemat blokowy układu pomiarowego charakterystyk anten metodą radio-astronomiczną

Zysk energetyczny anteny można obliczyć z następującej zależności [1]

$$G = (8\pi \cdot k \cdot T_a K_1 K_2) / S \lambda^2, \quad (1.7)$$

w której

$$K_1 = (a_0 l_0 + b_w l_w) \sec \Theta,$$

$$K_2 = \left( \iint_{\Omega_S} B(\Omega) d\Omega \right) \iint_{\Omega_S} B(\Omega) F(\Omega) d\Omega,$$

$a_0, b_w$  i  $l_0, l_w$  — odpowiednio wysokości i tłumienia warstw tlenu i pary wodnej w atmosferze,

$\Theta$  — kąt zenitalny ustawienia anteny,

$B(\Omega)$  — przestrzenny rozkład promieniowania gwiazdy,

$F(\Omega)$  — charakterystyka promieniowania anteny,

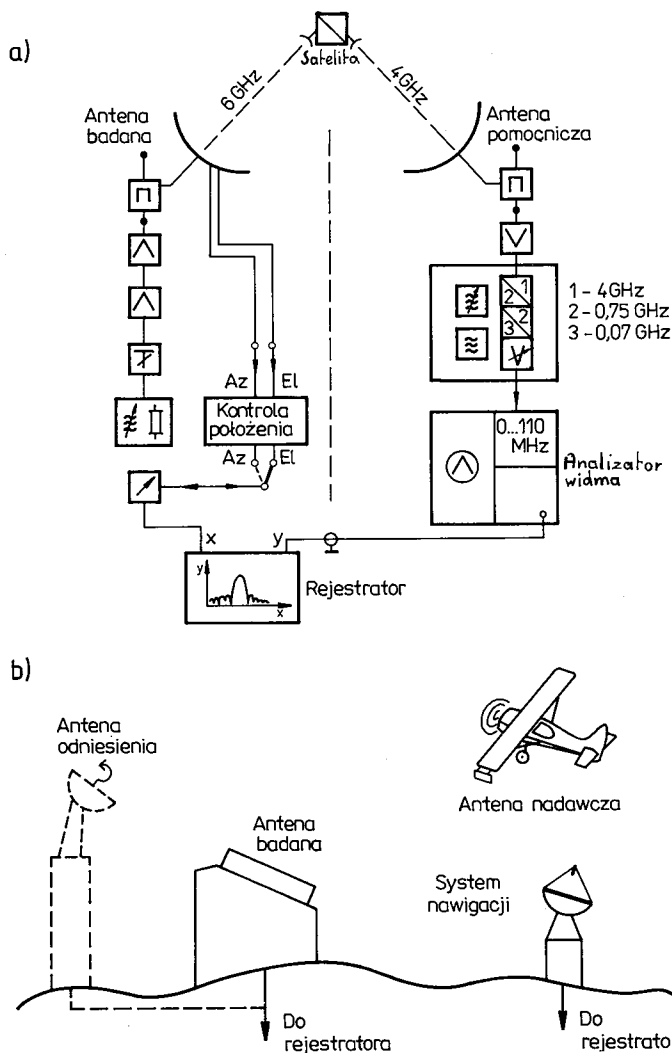
$S$  — gęstość mocy promieniowanej przez radio-gwiazdę.

Pomiar metodą radio-astronomiczną sprowadza się do określenia temperatury szumowej badanej anteny (rys. 7b)

$$T_a = T_k P_s / P_k. \quad (1.8)$$

#### Metody pomiarowe ze źródłami na samolotach

W przypadku gdy gęstość mocy radio-gwiazd jest zbyt mała dla celów pomiarowych (anteny o średnich kierunkowościach), korzysta się z nadajników umieszczonych na obiektach latających, poruszających się w strefie dalekiej badanej anteny (rys. 8).



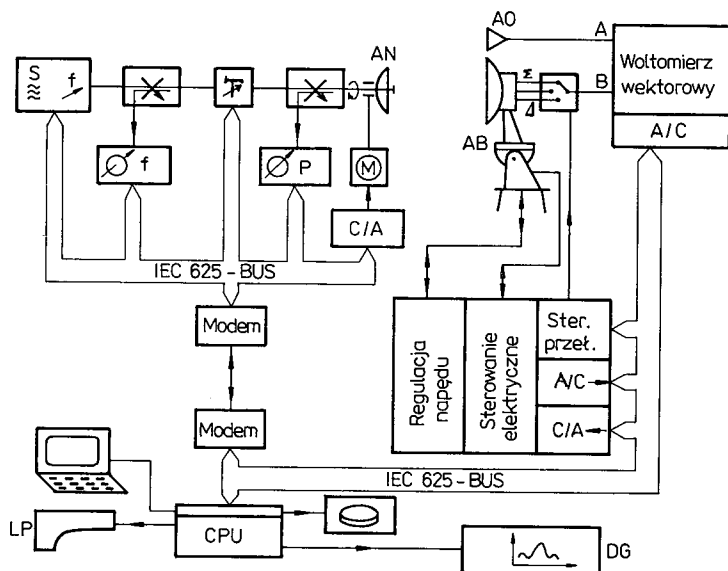
Rys. 8. Stanowiska do pomiaru charakterystyk anten z wykorzystaniem obiektów latających: Transpondera satelitowego (a) i nadajnika na samolocie (helikopterze) (b)

Zastosowanie satelitów komunikacyjnych jako odległych źródeł pola ma pewne zalety: satelita jest dostępny dla dużej liczby stacji naziemnych, badane anteny są mierzone przy kątach elewacji odpowiadających naturalnym warunkom pracy. W przypadku pomiarów samolotowych, system nawigacyjny (radio-lokacyjny, laserowy, optyczny) przekazuje informacje o położeniu źródła układowi rejestracji i przetwarzania danych.

#### 1.6. AUTOMATYCZNE STANOWISKO POMIAROWE

W ostatnich latach obserwuje się znaczne zainteresowanie automatycznymi systemami pomiarowymi, pozwalającymi na oszczędność czasu pracy ludzkiej,

zwiększenie szybkości pomiarów (np. 50 tys. pom./1,5 godz.), oraz wyeliminowanie błędów wnoszonych przez obsługę. Przykład stanowiska zautomatyzowanego przedstawiono na rys. 9, [5].



Rys. 9. Stanowisko do zautomatyzowanych pomiarów charakterystyk anten z wykorzystaniem komputera

## Część II

### 2. PODSTAWOWE PARAMETRY ANTEN

#### 2.1. CHARAKTERYSTYKA PROMIENIOWANIA

Każda składowa pola elektromagnetycznego wokół anteny może być przedstawiona w postaci ogólnej

$$\vec{E}_U = I_U E(\Theta, \Phi) \exp j\psi(\Theta, \Phi) \exp(-jkr)/r \quad (2.1)$$

gdzie:

$I_U$  — wektor jednostkowy, w ogólnym przypadku polaryzacji eliptycznej (zespolony),

$E(\Theta, \Phi)$  — rozkład amplitudy składowej elektrycznej pola promieniowanego,

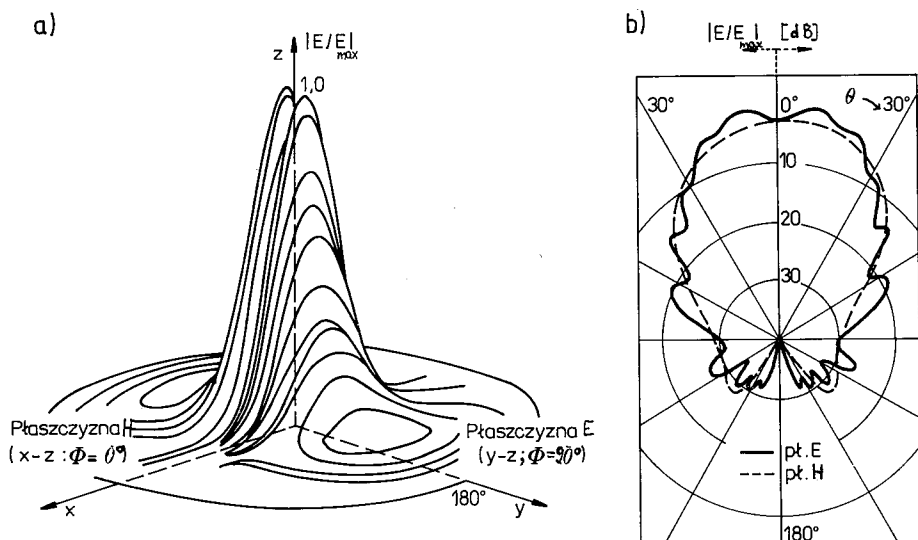
$\psi(\Theta, \Phi)$  — rozkład fazy pola promieniowanego,

$k$  — stała propagacji,

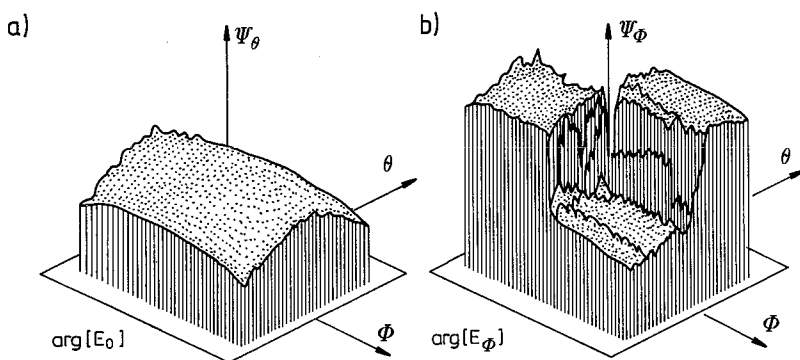
$r$  — odległość punktu pomiarowego od anteny.

W dużej odległości (strefa daleka — Fraunhofera) od anteny, składowe: elektryczna i magnetyczna pola promieniowanego przez antenę są ortogonalne względem siebie i kierunku propagacji, a stosunek ich amplitud jest stały i równy impedancji charakterystycznej ośrodka (pole ma strukturę fali płaskiej).

Charakterystykę promieniowania definiujemy jako rozkład natężenia pola elektrycznego na powierzchni kuli o promieniu  $R$  (2), której środek pokrywa się ze środkiem anteny [7]. Charakterystyka promieniowania przedstawia pewną zamkniętą powierzchnię, w ogólnym przypadku złożoną z kilku przestrzennych wiązek różnej postaci (rys. 10). Największą z nich nazywamy wiązką (listkiem) główną, pozostałe wiązkami (listkami) bocznymi. Wykonanie wykresu trójwymiarowego wymaga dużej ilości danych pomiarowych i jest kłopotliwe, dlatego ograniczamy się zwykle do podania dwóch wzajemnie prostopadłych przekrojów charakterystyki promieniowania w płaszczyźnie wektora pola elektrycznego ( $E$ ) i w płaszczyźnie wektora pola magnetycznego ( $H$ ). Przekroje przestrzennej charakterystyki promieniowania przedstawiamy w postaci wykresów w biegunowym (rys. 10b) lub prostokątnym układzie współrzędnych.



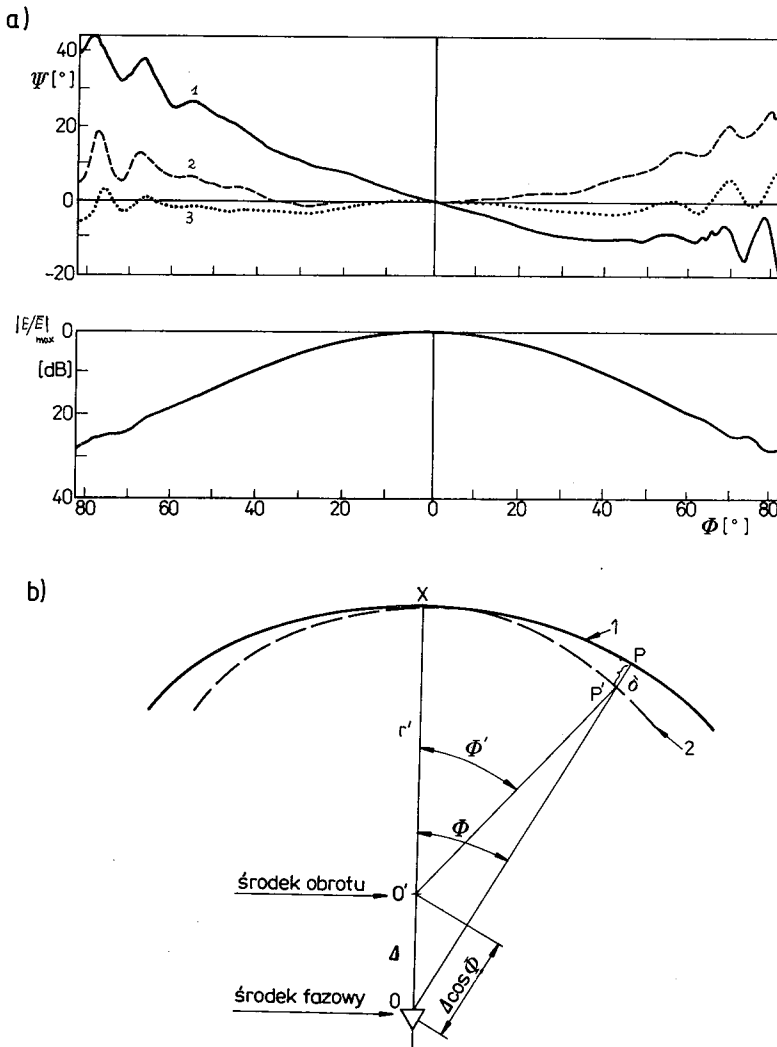
Rys. 10. Amplitudowe charakterystyki promieniowania anteny: w trójwymiarowym układzie współrzędnych (a) i jej przekroje w biegunowym układzie współrzędnych (b)



Rys. 11. Fazowe charakterystyki promieniowania anteny tubowej: dla składowej  $\Theta$  (a) i dla składowej  $\Phi$  (b) pola promieniowanego

Charakterystyki amplitudowe promieniowania pola nie dają informacji o fazie pola. W pewnych przypadkach (anteny fazowe, źródła oświetlające reflektory, itp.) znajomość fazy pola jest istotna i wtedy odwołujemy się do charakterystyk fazowych.

Charakterystyka fazowa (2.1) w kierunku wersora  $I_U$  jest wyrażona funkcją  $\psi(\Theta, \Phi)$ . Dla polaryzacji liniowej  $I_U$  jest rzeczywisty i może być reprezentowany przez  $I_\Theta$  i  $I_\Phi$ , odpowiednio w kierunkach  $\Theta$  i  $\Phi$ . Charakterystyka fazowa jest funkcją okresową z okresem  $2\pi$ . Dla wielu anten charakterystyka fazowa jest stała w obrębie jednego listka i zmienia się skokowo o kąt  $\pi$  przy przejściu od listka do listka. W ogólnym jednak przypadku charakterystyka fazowa może mieć skomplikowany kształt (rys. 11).



Rys. 12. Amplitudowe i fazowe charakterystyki promieniowania anteny (a) geometria ilustrująca odchylenie fazy od powierzchni ekwifazowej dla różnych położeni osi obrotu anteny (b) 1) sferyczne czoło fali ze środkiem w punkcie O 2) trasa przemieszczania sondy pomiarowej (środek obrotu w punkcie O')

Jeżeli powierzchnia ekwifazowa (czoło fali) ma kształt kuli to środek tej kuli nazywany jest środkiem fazowym anteny. Dla anten nie mających środka fazowego wprowadzamy pojęcie środka promieniowania, definiowanego jako środek kuli, której powierzchnia w pewnym zakresie kątów (zwykle w obrębie listka głównego) jest zbliżona do powierzchni ekwifazowej [7] (rys. 12a).

Odchylenie fazy czoła fali od sfery, w danym przedziale kąta bryłowego, zależy od typu i konstrukcji anteny oraz położenia środka sfery odniesienia.

Rozważmy sytuację przedstawioną na rys. 12b. [3]. Niech badana antena będzie umieszczona w punkcie 0, początku sferycznego układu współrzędnych, z główną wiązką skierowaną wzdłuż osi X. Natomiast środek obrotu anteny pomiarowej (sondy), niech będzie w punkcie 0'. Gdy sonda będzie położona w punkcie P (obrócona o kąt  $\Phi$  od osi anteny), wówczas słuszne jest równanie

$$r' + \Delta = \Delta \cos \Phi + r' + \delta, \quad (2.2)$$

ponieważ  $OX = OP$ .

Tak więc odchylenie  $\delta$  jest określone zależnością

$$\delta = (1 - \cos \Phi) \Delta. \quad (2.3)$$

Odchylenie fazy  $\alpha$  czoła fali od fazy sferycznego czoła fali odniesienia jest określone przez

$$\alpha = 2\pi\delta/\lambda [\text{rad}]. \quad (2.4)$$

Dokładność określenia zależy od szerokości przedziału kątowego  $\Phi$  w którym są wykonywane pomiary.

Określenie położenia środka fazowego polega na zmianie położenia  $\Delta$  środka obrotu sondy. Gdy  $\Delta = 0$ ,  $\delta$  musi być równe zero, przy założeniu kołowego czoła fali w danej płaszczyźnie. Faza mierzona przez sondę powinna być stała w danym przedziale kątów  $\Phi$ . W praktyce  $\delta$  nie jest równe zero w całym przedziale kątów  $\Phi$ , zatem odchylenie fazy  $\alpha$  ulega niewielkim zmianom (rys. 12).

## 2.2. POLARYZACJA ANTENY

Polaryzacja fali elektromagnetycznej jest definiowana jako właściwość opisująca czasowe zmiany kierunku i względnej amplitudy wypadkowego wektora pola elektrycznego  $E(r, t)$  [4]:

$$E(r, t) = I_\theta E_\theta(r, t) + I_\phi E_\phi(r, t), \quad (2.5)$$

gdzie:

$$E_\theta(r, t) = \text{Re}[E_\theta \exp(j\omega t - kr + \varphi_\theta)] = E_\theta \cos(\omega t - kr + \varphi_\theta) \quad (2.5a)$$

$$E_\phi(r, t) = \text{Re}[E_\phi \exp(j\omega t - kr + \varphi_\phi)] = E_\phi \cos(\omega t - kr + \varphi_\phi) \quad (2.5b)$$

$I_\theta, I_\phi$  — wektory kierunkowe.

Rugując we wzorze (2.5) członny zawierające  $(\omega t - kr + \varphi)$  otrzymujemy równanie elipsy polaryzacji

$$a(E_\theta^2 + E_\phi^2) - bE_\theta E_\phi = 1 \quad (2.6)$$

gdzie:

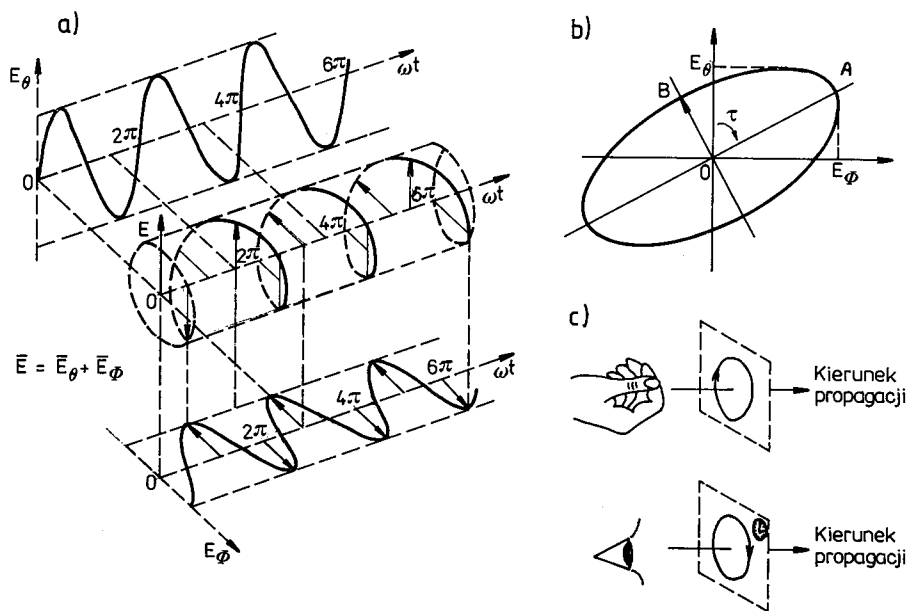
$$a = \sin^{-2} \Delta\varphi;$$

$$b = 2a \cos \Delta\varphi;$$

$$\Delta\varphi = \varphi_\phi - \varphi_\theta$$

$E_\theta, E_\phi$  — względne amplitudy składowych pola  $E$ .

Wektor ten jest sumą dwóch wektorów, wzajemnie prostopadłych, zmiennych okresowo w czasie. Wiruje więc on w przestrzeni i jednocześnie zmienia swą amplitudę, tak że jego koniec, w ogólnym przypadku opisuje elipsę (2.6) — mówimy, że pole jest spolaryzowane eliptycznie (rys. 13a).



Rys. 13. Wirujące w przestrzeni wektory składowe pola elektrycznego (a), parametry elipsy polaryzacji (b) i sposób określania skrętności elipsy polaryzacji fali (c)

Względne amplitudy i fazy składowych wektora  $E(r, t)$  zmieniają się wraz ze zmianą kątów  $\theta$  i  $\phi$ . Polaryzacja pola jest więc funkcją kierunku. Zależnie od kierunku obrotu wektora wypadkowego mówimy o polaryzacji lewoskrętnej [ $\Delta\varphi = -(0.5 + 2n)\pi$ ] (LH lub CCW) lub prawoskrętnej [ $\Delta\varphi = (0.5 + 2n)\pi$ ] (RH lub CW), gdzie  $n = 0, 1, 2, \dots$  (rys. 13c).

Szczególnymi przypadkami polaryzacji fali mogą być: polaryzacja kołowa [ $E_\theta = E_\phi$  i  $\Delta\varphi = \pm(0.5 + 2n)\pi$ ], liniowa ( $\Delta\varphi = n\pi$ ) lub eliptyczna [ $E_\theta \neq E_\phi$  i  $\Delta\varphi = \pm(0.5 + 2n)\pi$  lub  $\Delta\varphi = \pm 0.5 n\pi$ , niezależnie od wartości  $E_\theta$  i  $E_\phi$ ].

W przypadku eliptycznej polaryzacji fali określa się jej parametry (rys. 13b):

— współczynnik osiowy AR

$$AR = \text{wielka oś}/\text{mała oś} = OA/OB, \quad 1 \leq AR \leq \infty \quad (2.7)$$

gdzie:

$$OA_{(OB)} = \left( 0.5(E_\theta^2 + E_\phi^2 + \left[ E_\theta^4 + E_\phi^4 + 2E_\theta^2 E_\phi^2 \cos(2\Delta\varphi) \right]^{0.5}) \right)^{0.5} \quad (2.7a)$$

— nachylenie elipsy polaryzacji

$$\tau = 0.5\pi - 0.5 \arctg[2E_\theta E_\phi \cos(\Delta\varphi)/(E_\theta^2 - E_\phi^2)]. \quad (2.8)$$

Przez polaryzację anteny rozumiemy polaryzację fali elektromagnetycznej emitowanej przez nią (antena nadawcza).

### 2.3. KIERUNKOWOŚĆ I ZYSK ENERGETYCZNY

Kierunkowość (D) jest definiowana jako stosunek maksymalnej gęstości promieniowania (mocy promieniowanej w jednostkowym kącie bryłowym) do średniej gęstości promieniowania przez antenę. Kierunkowość anteny jest jednoznacznie określona przez jej charakterystykę promieniowania. W wielu spotykanych w praktyce przypadkach przestrzenna charakterystyka promieniowania może być przedstawiona w postaci iloczynu dwóch funkcji jednej zmiennej, wówczas wyrażenie na kierunkowość anteny możemy zapisać w postaci [7]:

$$D = 4\pi F^2(\Theta, \Phi) / \int_0^{2\pi} \int_0^\pi F^2(\Theta, \Phi) \sin \Theta d\Theta d\Phi = 4\pi / \int_0^{2\pi} F_1^2(\Phi) d\Phi \int_0^\pi F_2^2(\Theta) \sin \Theta d\Theta \quad (2.9)$$

gdzie:

$F_1(\cdot)$  i  $F_2(\cdot)$  — reprezentują charakterystyki promieniowania anteny, odpowiednio w płaszczyźnie  $\Phi$  i  $\Theta$ .

Dla anten o dużych kierunkowościach (z jednolistkową charakterystyką promieniowania), kierunkowość można wyliczyć, korzystając ze wzoru przybliżonego [4]:

$$D \cong 41\,253/\alpha_\theta \alpha_\phi \cong 72\,815/(\alpha_\theta^2 + \alpha_\phi^2) \quad (2.10)$$

gdzie:

$\alpha_\theta$  i  $\alpha_\phi$  szerokości wiązki głównej (w stopniach), określane dla spadku poziomu o -3 dB od wartości maksymalnej

Kierunkowość anteny nie uwzględnia strat mocy w antenie. Wprowadza się wielkość określającą właściwości kierunkowe anteny z uwzględnieniem jej sprawności — zysk energetyczny (G).

Zysk G definiujemy jako stosunek maksymalnej gęstości promieniowania wytwarzanej przez antenę do całkowitej mocy doprowadzonej do jej zacisków lub maksymalnej gęstości promieniowania wytwarzanej przez antenę wzocową zasilaną tą samą mocą [7]:

$$G_B = 4\pi U_B/P_B = (P_B(\text{ant. bad.})/P_0(\text{ant. odn.}))G_0, \quad (2.11)$$

gdzie:

$P_B, P_0$  — moce na zaciskach anten;

$U_B$  — gęstość promieniowania;

$G_B, G_0$  — zyski.

Za antenę wzorcową (odniesienia), zwykle, w zakresie mikrofal przyjmuje się antenę tubową lub bezstratną antenę izotropową (rzadziej — dipol półfalowy).

Między kierunkowością a zyskiem energetycznym anteny zachodzi prosta zależność

$$G = \eta D \quad (2.12)$$

gdzie:

$\eta = \eta_r \eta_p \eta_m \eta_d$  — sprawność anteny,

$\eta_r = (1 - |\Gamma|^2)$  — sprawność dopasowania,

$\Gamma = (Z_{we} - Z_o)/(Z_{we} + Z_o)$  — współczynnik odbicia na wejściu anteny,

$Z_{we}$ ,  $Z_o$  impedancje: wejściowa anteny i charakterystyczna toru zasilającego,

$\eta_p = |\mathbf{1}_o \mathbf{1}_p^*|^2 = |\cos \xi|^2$  — sprawność dopasowania polaryzacyjnego

$\mathbf{1}_o$ ,  $\mathbf{1}_p$  — wersory kierunkowe polaryzacji anteny i fali padającej,

$\xi$  — kąt zawarty pomiędzy wersorami,

$\eta_m$ ,  $\eta_d$  — straty w przewodniku i dielektryku anteny.

Dla wielu anten zysk energetyczny można określić z zależności przybliżonej

$$G \cong 30\,000/\alpha_\phi \alpha_\theta \quad (2.13)$$

## Część III

### 3. POMIARY PODSTAWOWYCH PARAMETRÓW ANTEN

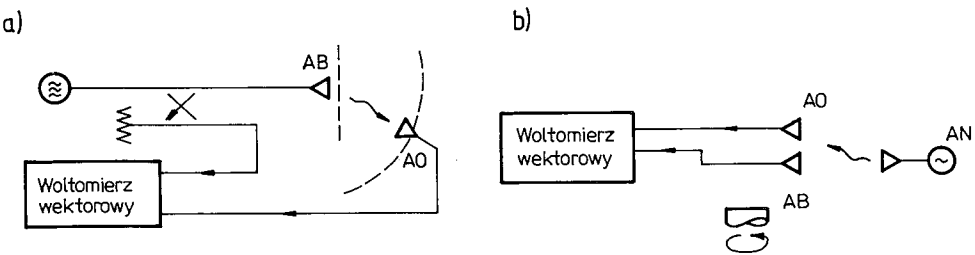
#### 3.1. POMIARY AMPLITUDOWYCH CHARAKTERYSTYK PROMIENIOWANIA

Antena badana pracuje zwykle w charakterze anteny odbiorczej (jedynie wówczas gdy zawiera elementy nieodwracalne (nieliniowe) jest badana zgodnie z przeznaczeniem). Charakterystyka jest mierzona przez obrót anteny badanej, umieszczonej na stole obrotowym, w obu płaszczyznach i przedstawiana w postaci wykresu w wybranym układzie współrzędnych (ploter) lub rejestrowane. Pomiary są wykonywane w układzie jak na rys. 2.

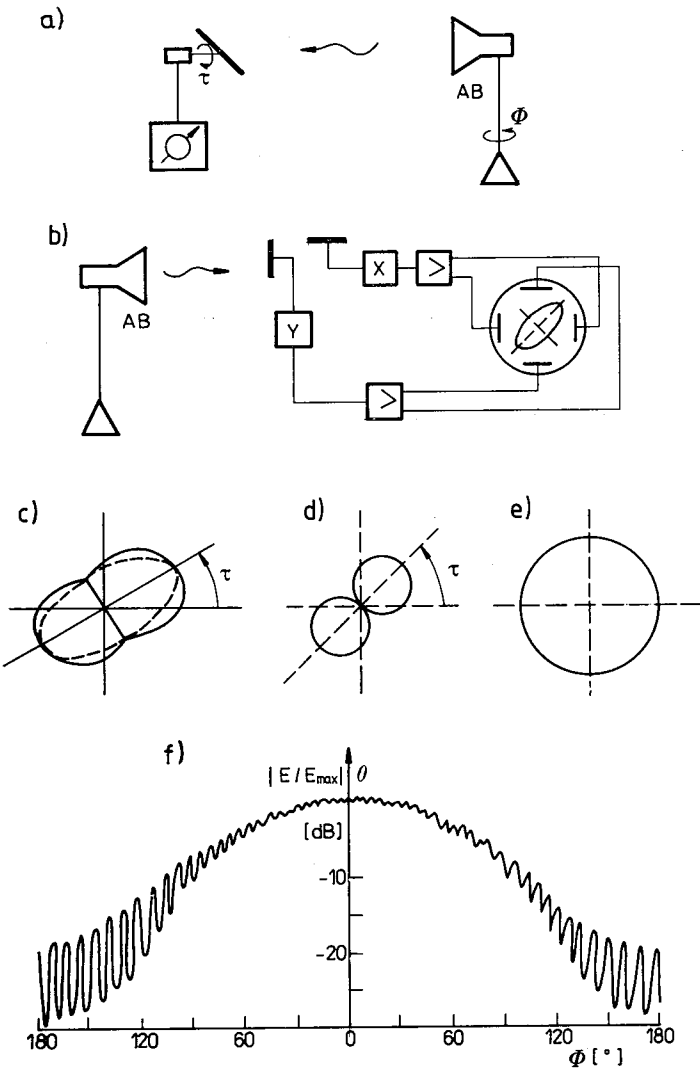
#### 3.2. POMIARY FAZOWYCH CHARAKTERYSTYK PROMIENIOWANIA

Charakterystyka fazowa  $\psi(\Theta, \Phi)$  anteny jest funkcją wartości okresowej z okresem  $(2\pi)$  i jest wielkością względną.

Podstawowe dwa układy pomiaru charakterystyk fazowych dla małych i dużych odległości między antenami przedstawiono na rys. 14. Pomiary charakterystyk fazowych, podobnie jak w amplitudowych wykonuje się dla stałej odległości (strefa daleka) między antenami. W przypadku małej odległości między antenami, antena badana jest nieruchoma i pracuje jako nadawcza (rys. 14a), a ruchoma sonda pomiarowa porusza się po powierzchni sfery, w której środku znajduje się antena badana. Dla dużych odległości — antena badana jest umieszczona na stole obrotowym (pracuje jako — odbiorcza), przy czym sygnał odbierany przez nią jest porównywany z sygnałem innej nieruchomej anteny odniesienia w mierniku fazy (rys. 14b).



Rys. 14. Układy pomiarowe charakterystyk fazowych anten: dla małych (a) i dużych (b) odległości między antenami



Rys. 15. Układy pomiarowe dla wyznaczania własności polaryzacyjnych anten: metodą charakterystyki polaryzacyjnej (a) metodą amplitudową (b) i charakterystyki polaryzacyjne (c), (d), (e), (f)

W trakcie pomiarów należy zachować stałą długość drogi elektrycznej dla sygnału badanego i odniesienia (zmiany mogą występować podczas zginania lub skręcania kabli połączeniowych w.c.z., zmian temperatury i wilgotności otoczenia, oraz w wyniku odbić sygnałów na niedopasowanych połączeniach, itp.).

W celu wyeliminowania wpływu fluktuacji amplitudy sygnału odbieranego przez anteny stosuje się, przed porównaniem sygnałów, modulację sygnału w jednym z kanałów pomiarowych i pomiar fazy między nimi za pomocą odbiornika homodynowego.

### 3.3 POMIAR POLARYZACJI ANTENY

Pełny opis właściwości polaryzacyjnych anteny wymaga znajomości elipsy polaryzacji: współczynnik osiowy, kąt nachylenia, skrętność (kierunek wirowania wektora pola  $E$ ). Najprostszą, jest metoda pomiaru charakterystyki polaryzacji, która określa kształt elipsy, bez skrętności (rys. 15a). Przy pomiarach metodą charakterystyki polaryzacji, antena badana jest nieruchoma i pracuje jako nadawcza. W odległym miejscu jest ustawiona sonda (o polaryzacji liniowej, np. dipol, tuba) obracana w płaszczyźnie polaryzacji. Typowe charakterystyki przedstawiono na rys. 15c, d, e, f. W metodzie amplitudowej (rys. 15b) wykorzystuje się dwie ortogonalne anteny spolaryzowane liniowo, umieszczone ortogonalnie w płaszczyźnie polaryzacji.

W celu określenia skrętności polaryzacji anteny, pomierzonej jedną z powyższych metod, należy wykonać dodatkowy pomiar, np. stosując dwie pomiarowe anteny spolaryzowane kołowo o przeciwnych skrętnościach. Skrętność polaryzacji badanej odpowiada skrętności polaryzacji anteny, na której zaciskach wyindukuje się większe napięcie.

### 3.4. POMIAR ZYSKU ENERGETYCZNEGO

Najczęściej, w praktyce, stosuje się dwie metody pomiaru zysku: pomiar zysku bezwzględnego (odniesionego do bezstratnej anteny izotropowej) i pomiar zysku odniesionego do anteny wzorcowej (metoda porównawcza). Pomiar zysku bezwzględnego stosuje się dla wyznaczenia zysku anten wzorcowych, wykorzystywanych do pomiarów zysków metodą porównawczą.

Metoda zysku bezwzględnego opiera się na równaniu Friis'a:

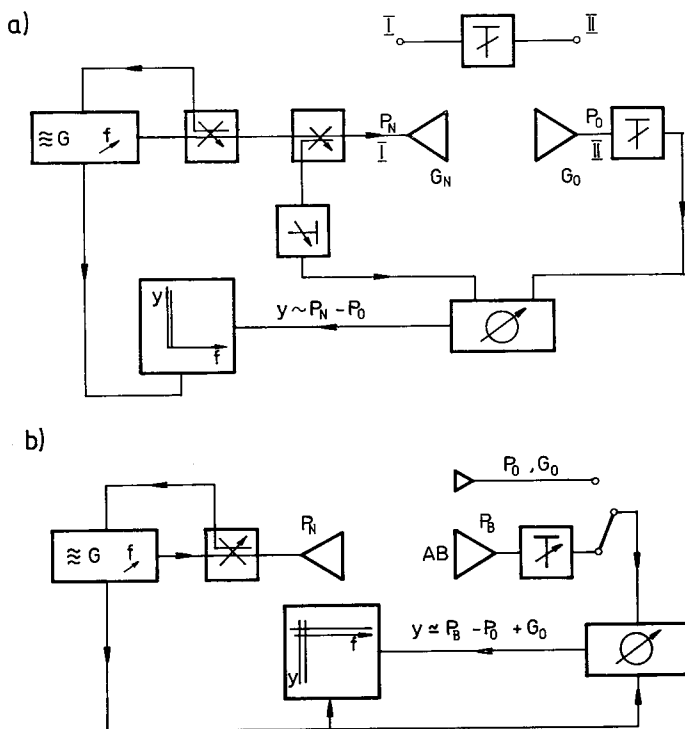
$$G_N + G_0 = 20 \log(4\pi R/\lambda) + 10 \log(P_O/P_N), \quad (3.1)$$

gdzie:

$G_N, G_0, P_N, P_O$  — zyski i moce (na zaciskach), odpowiednio, anten: nadawczej i odbiorczej,

$R$  — odległość między antenami.

Zakłada się przy tym, że anteny są dopasowane polaryzacyjnie i zorientowane na kierunek maksymalnego promieniowania (w strefie dalekiej). Gdy obie anteny są identyczne, dla ustalonych  $R$  i  $\lambda$ , wystarczy pomierzyć  $(P_O/P_N)$ .



Rys. 16. Schematy blokowe układów do pomiaru zysku energetycznego anten: metodą — zysku bezwzględnego (a) i metodą porównawczą (b)

Często, w celu uniknięcia wpływu nieliniowości odbiornika (detektora) w funkcji poziomu mocy, pomiar zysku wykonuje się przy stałym poziomie mocy odbieranej. W tym celu, po określeniu poziomu sygnału na wejściu odbiornika przy transmisji przez anteny zamiast anten: odbiorczej i nadawczej włącza się precyzyjny tłumik regulowany 16a. Dla stałych mocy  $P_N$  i  $P_0$ , przy jednakowych antenach, po odczytaniu tłumienia  $L_0$  (odpowiadającego poziomowi mocy przy transmisji przez anteny) i  $L_R$  (dla bezpośredniego połączenia nadajnika z odbiornikiem) tłumika, możemy określić zysk energetyczny anteny, korzystając ze wzoru

$$G = [20 \lg(4\pi R/\lambda) - (L_0 - L_R)]/2. \quad (3.2)$$

W przypadku, gdy anteny nie są identyczne, należy zastosować trzecią antenę, oraz dla warunków jak wyżej, określić stosunki mocy w trzech kombinacjach zastosowania anten (każda z każdą) w układzie pomiarowym jak na rys. 16a.

W porównawczej metodzie pomiarowej (rys. 16b) wykorzystujące się anteny wzorcowe (rezonansowy dipol półfalowy o zysku  $G = 2.15$  dB lub standardową antenę tubową o zysku  $G = 12 \div 25$  dB) o znanym zysku. Obydwie anteny mają polaryzację liniową. Równanie Friis'a upraszcza się w tym przypadku do postaci

$$G_B = G_0 + 10 \lg(P_B/P_0). \quad (3.3)$$

W przypadku, gdy badana antena jest spolaryzowana eliptycznie (kołowo) jako antenę wzorcową (odniesienia) stosuje się antenę o identycznej polaryzacji lub dokonuje się pomiaru w dwóch płaszczyznach: pionowej  $G_{0V}$  i poziomej  $G_{0H}$ , wówczas zysk określa się z zależności

$$G_B = G_{0V} + G_{0H}. \quad (3.4)$$

#### BIBLIOGRAFIA

1. IEEE *Standard Test Procedure for Antennas* IEEE Std. 149-1979, published by IEEE Inc., 1979
2. W.H. Kummer, E.S. Gillespie: *Antenna Measurements* - 1978. Proc. IEEE, vol. 66, No. 4, pp. 483-507, April 1978.
3. *Basic Antenna Measurements*. Scientific - Atlanta, Inc.
4. C.A. Balanis: *Antenna Theory Analysis and Design*. Harper & Row Publishers, New York, 1982
5. E. Stirner: *Antennen. Band 3: Messtechnik*. Heidelberg: Huthing, 1985
6. E.B. Joy, W. Marshall Leach, G.P. Radrique, D.T. Paris: *Applications of Probe-Compensated Near-Field Measurements*. IEEE Trans., vol. AP-26, No. 3, pp. 379-89, May, 1978
7. D.J. Bem: *Anteny i rozchodzenie się fal radiowych*. Wydawnictwa Naukowo-Techniczne, Warszawa, 1973

W.J. KRZYSZTOFIK

#### MODERN MEASURING METHODS OF MICROWAVE ANTENNAS

##### Summary

The rapid development of communication systems on sea and on land (in particular satellite communication) renders the determination of the main parameters of applied antennas necessary. The compatibility requirements set before different radio communication systems, applied by different services, become critical. This is to be observed, in particular on the limited space of ships. The first part defines the conditions to be met by a test stand and presents the substantial types of these stands (in close, intermediary and remote fields). The second part defines the so-called field characteristics (complex radiation pattern, polarization, directivity, power gain) of antennas and block diagrams of measuring systems are presented.

# Badania symulacyjne wymiany informacji między dwoma węzłami magistralowej lokalnej sieci przetwarzania danych

KRZYSZTOF ANTOSIK, JERZY HOSZA, TADEUSZ STACHOWIAK

*Instytut Podstaw Informatyki, Polska Akademia Nauk*

*Otrzymano 1989.05.20*

W artykule przedstawiono symulacyjne badanie wymiany informacji między dwoma węzłami w lokalnej sieci magistralowej z dostępem rywalizacyjnym typu CSMA-CD. Stanowi to pewien etap badań prowadzonych w Instytucie Podstaw Informatyki PAN nad możliwością zastosowania lokalnych sieci przetwarzania danych dla potrzeb medycyny. W pierwszej części pracy omówiono krótko magistralową sieć lokalną będącą przedmiotem rozważań. Następnie przedstawiono opis modelu symulacyjnego, przedyskutowano ważniejsze problemy w realizacji programowej oraz omówiono sposób weryfikacji wyników. W końcowej części opisano przeprowadzone eksperymenty symulacyjne i skomentowano uzyskane wyniki.

## 1. WPROWADZENIE

W ostatnich latach szeroko rozwijają się lokalne sieci komputerowe służące celom przetwarzania danych. Znajdują one zastosowanie wszędzie tam, gdzie istnieje potrzeba wymiany informacji między urządzeniami na niewielkim obszarze obejmującym swym zasięgiem teren szpitala, zakładu przemysłowego, ośrodka akademickiego czy urzędu. Za pomocą sieci lokalnych realizowane jest zarządzanie, zbieranie danych, przesyłanie wiadomości, rozdział funkcji obliczeniowych — a więc to co najpełniej pozwala wykorzystać rozproszony charakter zadań użytkowych. Sterowanie przepływem informacji jest osiąganе za pomocą zestawu reguł tworzących protokoły komunikacyjne.

Jedną z dziedzin, w której sieci lokalne znajdują zastosowanie jest medycyna. W krajach o zaawansowanej technice istnieją systemy informatyczne w szpitalach, ambulatoriach, opiece pielęgniarstwie [6], [9]. Wykorzystuje się kliniczne wspomaganie procesu leczenia [12]. Prowadzi się różnorodne analizy danych biomedycznych [11]. Mikrokomputery są stosowane w ultrasonografii, medycynie nuklearnej, radiologii, analizie EKG, epidemiologii i statystyce medycznej, analizie EEG, laboratorium analitycznym, intensywnej terapii, farmacji i gospodarce lekami [6], [7]. W Polsce prace w tym zakresie prowadzi część placówek służby zdrowia, ale w większości są to systemy specjalistyczne, uwzględniające potrzeby lokalne. W Politechnice Wrocławskiej prowadzone są prace, które mają doprowadzić do określenia

wymagań stawianych lokalnym sieciom komputerowym budowanym dla potrzeb szpitalnictwa [7]. Zgodnie z przedstawioną koncepcją system szpitalny ma być wyspecjalizowanym systemem wielodostępnym o strukturze rozproszonej. Istotną zaletą tej struktury jest podział funkcji systemu pomiędzy wiele komputerów oraz możliwość rozproszenia bazy danych. Dzięki temu uzyskuje się na ogół skrócenie czasów reakcji systemu, szczególnie przy realizacji funkcji lokalnych, tzn. funkcji wykorzystywanych w obrębie jednego komputera.

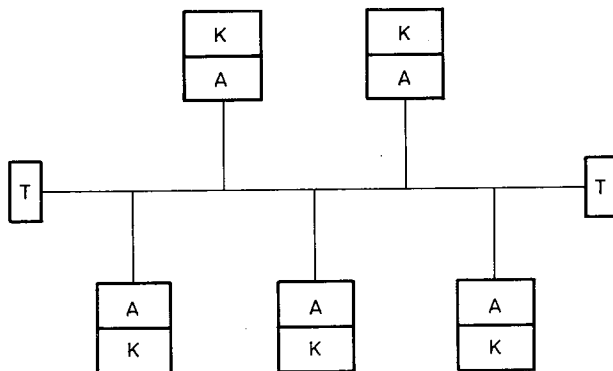
Sprawne działanie informatycznego systemu szpitalnego jest ściśle związane z koniecznością wymiany informacji między jego poszczególnymi obiektami oraz sterowaniem rozproszonych jednostkowych procesów i pojedynczych operacji. Pojawia się więc potrzeba systemu komunikacji, który zapewni realizację wymienionych funkcji. Problem ten można rozwiązać właśnie przy pomocy lokalnych sieci komputerowych [4]. Jednak przed zainstalowaniem takiej sieci należy, w procesie projektowym, sprawdzić różne założenia i warianty. Można to zrobić stosunkowo tanim kosztem przy pomocy badań symulacyjnych. Stwarza to duże możliwości do przeprowadzenia analizy pracy systemu sterowania, jakim jest sieć lokalna służąca celom przetwarzania [2]. Takie podejście pozwala na optymalizację trudną do osiągnięcia innymi metodami [13], [15].

W pracy przedstawiono symulacyjne badanie wymiany informacji między dwoma węzłami w lokalnej sieci magistralowej z dostępem rywalizacyjnym typu CSMA-CD (Carrier Sense Multiple Access with Collision Detection). Sieci tego typu są w trakcie budowania w kraju [1], [4], [14], jednak możliwości eksperymentowania w nich są ograniczone. Takich ograniczeń nie ma podczas symulacji komputerowej, do której sieć fizyczna nie jest konieczna. Podjęte badania stanowią pierwszy krok na drodze tworzenia modelu symulacyjnego dla całej sieci lokalnej. W pierwszej części pracy omówiono krótko magistralową sieć lokalną będącą przedmiotem rozważań. Następnie przedstawiono opis modelu symulacyjnego, przedyskutowano ważniejsze problemy w realizacji programowej oraz omówiono sposób weryfikacji wyników. Dalej opisano przeprowadzone eksperymenty symulacyjne i skomentowano uzyskane wyniki oraz zamieszczono uwagi końcowe.

## 2. MAGISTRALOWA SIEĆ LOKALNA Z DOSTĘPEM RYWALIZACYJNYM

Typową strukturę magistralowej sieci lokalnej przedstawiono na rys. 1. Wszystkie stacje (np. mikrokomputery) są podłączone za pomocą adapterów do jednego wspólnego ośrodka transmisyjnego. Informacje wysyłane przez jedną ze stacji rozchodzą się do wszystkich pozostałych. Stacja, która w odbieranej wiadomości rozpozna swój adres przystępuje do przetwarzania informacji. W przeciwnym przypadku informacja jest przez nią pomijana. Ośrodek transmisyjny jest fizycznym nośnikiem informacji. Do budowy sieci lokalnych zwykle stosuje się a) kable symetryczne, b) kable koncentryczne, c) światłowody.

Kabel symetryczny jest bardzo rozpowszechnionym środkiem transmisyjnym. Nadaje się do transmisji rzędu kilku Mb/s. Przyjęcie takiego rozwiązania odznacza się



Rys. 1. Struktura magistralowej sieci lokalnej; A — adapter, K — urządzenie końcowe, komputer, T — terminator

małym kosztem i może być z powodzeniem stosowane w przypadku instalacji w pojedynczym budynku. Kabel symetryczny jest jednak mało odporny na wszelkiego rodzaju przeniki. Odpowiednie ekranowanie może częściowo zredukować to szkodliwe zjawisko. Większą szybkość przesyłania danych oraz większy zasięg zapewnia kabel koncentryczny. Jest on ponadto bardziej odporny na zakłócenia i interferencje. Na kablu współosiowym mogą być stosowane dwie metody transmisyjne: transmisja w pasmie podstawowym (baseband) i transmisja szerokopasmowa (broadband). Zasadnicza różnica między nimi polega na tym, że w pierwszym przypadku przesyła się jeden strumień danych, natomiast w drugim następuje jednocześnie przesyłanie w wielu kanałach. Rozwiązaniem przyszłościowym w sieciach lokalnych jest światłowód. Zastosowanie światłowodów w sieciach magistralowych jest jednak ograniczone, gdyż każde złącze wprowadza duże straty mocy i jest przyczyną optycznych odbić. W sieciach lokalnych o innych strukturach topologicznych zagadnienie to przedstawia się inaczej.

Jak już zaznaczono, ważnym elementem sieci lokalnej są adaptery. Służą one do połączenia urządzenia końcowego typu terminal, komputer czy drukarka do magistrali. Spełniają one przy tym następujące funkcje:

- realizacja zadań komunikacyjnych związanych z nadawaniem i odbieraniem sygnału z linii,
- realizacja dostępu do ośrodka,
- realizacja zadań komunikacyjnych związanych z wymianą informacji między dwoma punktami (adapterami) sieci,
- wymiana danych między adapterem a stacją.

W ogólnym przypadku jeden adapter może przyłączyć do sieci kilka urządzeń końcowych. Taki adapter jest wówczas nazywany grupowym (wieloportowym).

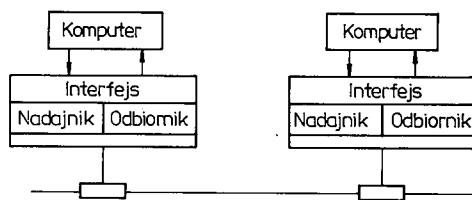
Sieć magistralowa z dostępem rywalizacyjnym należy do klasy sieci o sterowaniu zdecentralizowanym. Podstawą pracy takiej sieci jest metoda testowania zajętości łącza. Każda stacja, która chce nadawać prowadzi stały nasłuch. Jeżeli stwierdzi, że łącze jest wolne, rozpoczyna nadawanie. Można zauważyć, że metoda ta nie chroni przed wystąpieniem kolizji, czyli sytuacji, w której więcej niż jedna stacja nadaje w tym

samym czasie i ich sygnały zagłuszają się wzajemnie. Wynika to po pierwsze, z opóźnień pomiędzy momentem wykrycia, że łącze jest wolne a momentem rozpoczęcia nadawania, po drugie, z niezerowego prawdopodobieństwa, że co najmniej dwie stacje dokonują badania łącza i stwierdzają, że jest ono wolne w tej samej chwili. Aby temu przeciwdziałać, wprowadza się mechanizm wykrywania kolizji. Polega to na wprowadzaniu nasłuchu w czasie nadawania. Jeżeli nadająca stacja stwierdza, że nastąpiło zakłócenie transmisji nadawanie zostaje przerwane. Wysyła się krótki sygnał alarmowy (jam signal) informujący, że bieżąca transmisja jest nieudana. Po upływie pewnego, zwykle wyznaczonego w sposób losowy, odcinka czasu stacja ponawia sprawdzanie stanu łącza i podejmuje próbę retransmisji.

Opisana metoda dostępu do łącza, zwana CSMA-CD, jest bardzo rozpowszechniona, między innymi w sieciach typu Ethernet. Została również przyjęta jako standard IEEE 802.3.

### 3. OPIS MODELU SYMULACYJNEGO

Podstawowy schemat modelowanej sieci przedstawiono na rys. 2. Komputery dołączone są do ośrodka transmisyjnego za pomocą adapterów. Połączenie komputer-adapter ma charakter dwukierunkowy. Oznacza to, że jednocześnie może odbywać



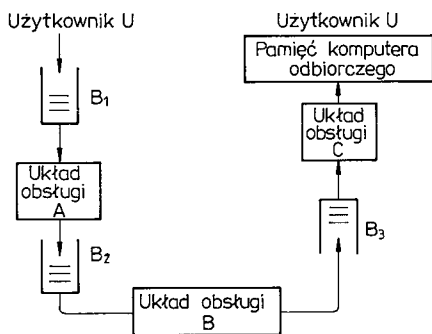
Rys. 2. Podstawowy schemat modelowanej sieci

się transmisja w obu kierunkach. Należy podkreślić, że taka sytuacja nie zawsze jest możliwa. Zależy to po prostu od budowy adaptera i sposobu jego połączenia z komputerem. Natomiast połączenie adapter-magistrala ma charakter półdwukierunkowy. W związku z tym w danej chwili może odbywać się transmisja tylko w jednym kierunku. Wynika to z charakteru rywalizacyjnej metody dostępu do ośrodka.

Współpraca między komputerem a adapterem przebiega następująco. Utworzona w komputerze wiadomość zostaje podzielona na bloki o długości mniejszej lub równej ustalonej dla danej sieci lokalnej. Bloki są numerowane według kolejności przeznaczonej do wysłania i umieszczane w kolejce do nadawania. Po stwierdzeniu, że adapter jest gotowy do odbioru następuje przesłanie do adaptera. W adapterze blok zostaje obudowany tworząc ramkę. Polega to na dodaniu adresów docelowych i źródłowych, preambuły (siedmiobajtowego pola służącego do synchronizacji układów odbiornika), informacji sterujących i kontrolnych. Gotowe ramki są ustawiane w kolejce do wysłania w magistralę, którą może stanowić np. kabel współosiowy. Dostęp do ośrodka transmisyjnego odbywa się zgodnie z omówionym w p. 2 protokołem CSMA-CD.

Należy zwrócić uwagę na pewnego rodzaju „podwójne” adresowanie wysyłanej ramki. Początkowo dołączany jest adres punktu dostępu do usług, czyli adres komputera źródłowego i docelowego. Informacja ta jest istotna, gdyż może być kilka lub więcej punktów udostępniania usług (tzw. portów) w adapterze. Dotyczy to sytuacji, gdy w sieci lokalnej jest adapter grupowy, do którego podłączono kilka urządzeń końcowych (np. różnych terminali lub komputerów). Natomiast adresy dołączane na poziomie dostępu do magistrali określają po prostu adapter źródłowy i adapter komputera, dla którego jest przeznaczona dana ramka.

Po odbiorze ramki z magistrali następuje przetworzenie informacji kontrolno-sterujących. Wydzielony z ramki blok jest umieszczany w kolejce wyjściowej do właściwego komputera (terminala). Po przydzieleniu przez ten komputer miejsca w pamięci następuje przesłanie bloku. W komputerze odbywa się proces składania wiadomości z bloków przesłanych przez adapter. Ideę takiego postępowania pokazano na rys. 3. Bufor  $B_1$  symbolizuje kolejkę zgłoszeń (wiadomości) podzielonych na bloki w komputerze użytkownika  $U_1$ . Układ obsługi  $A$  realizuje transmisję bloków z



Rys. 3. Model kolejkowy przesyłu informacji od użytkownika  $U_1$  do użytkownika  $U_2$

komputera do adaptera i proces formowania ramek. Gotowe do wysłania ramki są ustawiane w kolejce w buforze  $B_2$ . W układzie obsługi B uwzględniono proces dostępu do ośrodka, czas transmisji w magistrali oraz dekodowanie ramki w adapterze stacji odbiorczej. Utworzone w ten sposób bloki są umieszczane w buforze  $B_3$  adaptera odbiorczego. Czas związany z transmisją na trasie adapter — komputer oraz proces składania wiadomości z bloków uwzględniono w układzie obsługi C. Oczywiście dla transferu informacji od użytkownika  $U_2$  do użytkownika  $U_1$  obowiązuje analogiczny model kolejkowy.

#### 4. PROBLEMY REALIZACJI PROGRAMOWEJ MODELU

Z dotychczasowego opisu wynika, że symulacja sieci lokalnej polega na śledzeniu zmian stanu tej sieci w czasie. Można przyjąć, że dominują zmiany mające charakter nieciągły. Wówczas taki system nazywamy dyskretnym [3]. Ważną rolę w symulacji

systemów dyskretnych odgrywa pojęcie zdarzenia. Określamy je jako zmianę stanu w modelu w określonej chwili. Zdarzenie może spowodować zmianę wartości parametrów obiektu, utworzyć lub zlikwidować obiekt, uruchomić lub wstrzymać działanie. Symulacja polega na obserwacji zdarzeń czyli zmian stanów.

Istotnym zagadnieniem jest sterowanie procesem symulacyjnym, co wiąże się z realizacją upływu czasu. Istnieją dwie podstawowe metody aktualizacji czasu symulacyjnego. Pierwsza z nich polega na zwiększaniu tego czasu o stałą wartość i sprawdzaniu co się zmieniło w modelu. Jest to metoda stałego kroku. Drugi sposób polega na ustawianiu zegara na chwilę, w której ma wystąpić kolejne zdarzenie. Jest to metoda kolejnych zdarzeń. I właśnie ta metoda została wybrana przy tworzeniu programu symulacyjnego.

W metodzie kolejnych zdarzeń może zaistnieć taka sytuacja, że w w danej chwili czasu symulacyjnego zajdzie kilka zdarzeń jednocześnie, np. rozpoczęcie transmisji przez komputer i adapter. Są to tzw. zdarzenia równoczesne. Ponieważ komputer działa szeregowo a nie równolegle, więc nie jest możliwe bezpośrednie odtworzenie takich zdarzeń. Muszą one być realizowane w określonej kolejności. Istnieje więc potrzeba wprowadzenia dodatkowego mechanizmu, który by mógł rozstrzygać w tego typu sytuacjach. W przyjętym rozwiązaniu zaproponowano losowy dobór kolejności, w sytuacjach gdy nie ma innych przesłanek, np. natury praktycznej.

W realizacji programowej modelu wyróżnia się dwie zasadnicze fazy. Faza pierwsza dotyczy tzw. rozruchu programu. Zadaniem tego etapu jest eliminacja warunków początkowych, czyli przejście od obciążenia zerowego do nominalnego. Nie ma prostych reguł określających długość tego przedziału. Ten fragment programu jest realizowany następująco.

Oblicza się średnie opóźnienie bloku (ramki) w regularnych odstępach czasu  $\alpha$  i porównuje ze sobą. Jeżeli błąd względny między dwiema kolejnymi wartościami jest mniejszy od ustalonej wartości  $\delta$ , to przyjmuje się, że stan normalnej pracy został osiągnięty i program przechodzi do właściwej symulacji czyli fazy głównej. Jeżeli po kilkunastu próbach ten warunek nie jest spełniony, to przerywa się wykonywanie programu i wydłuża okres niezbędny do właściwego określenia średniego opóźnienia. W przeprowadzonych eksperymentach przyjęto, że błąd względny nie może przekraczać 5%. Natomiast właściwy odstęp czasu  $\alpha$  odpowiadał zwykle obsłudze kilku tysięcy bloków. Etap rozruchu nie jest brany pod uwagę przy końcowym obliczaniu wyników symulacji.

W fazie głównej odbywa się najważniejsza część symulacji działania sieci. Zmiana stanu w modelu symulacyjnym oznacza przejście między dwoma kolejnymi „istotnymi zdarzeniami” w programie. Lista tych istotnych zdarzeń, niezbędnych do właściwego odzwierciedlenia działania sieci lokalnej, jest następująca:

- a) powstanie nowej wiadomości,
- b) początek transmisji bloku w kanale komputer-adapter,
- c) koniec transmisji bloku w kanale komputer-adapter,
- d) początek transmisji w magistrali,
- e) kolizja w magistrali,
- f) koniec transmisji w magistrali.

Istotnym elementem programu jest procedura generowania zgłoszeń. Ponieważ program modeluje przepływ informacji między dwoma komputerami (stacjami), wiadomości powstają zarówno w jednej, jak i w drugiej stacji. W tym modelu zakłada się, że źródła ruchu są poissonowskie, co oznacza, że odstępy czasowe pomiędzy kolejnymi wiadomościami mają rozkład wykładniczy z zadaniem parametrem  $\frac{1}{\lambda}$ . Jego

odwrotność czyli  $\lambda$  jest intensywnością powstawania nowych wiadomości. Nie jest to jedyny element losowy w tym modelu. Długość wiadomości również jest wielkością stochastyczną. Przyjęto, że ma ona rozkład geometryczny o zadanej wartości średniej.

Przyjęty model zgłoszeń nie stanowi żadnego ograniczenia w możliwościach wykorzystania omawianego programu symulacyjnego. Strumień wiadomości dobiera się w zależności od istniejącej potrzeby. Można więc uwzględnić specyfikę urządzeń pracujących w informatycznym systemie szpitalnym.

## 5. WERYFIKACJA WYNIKÓW

Dokładna weryfikacja otrzymywanych wyników jest możliwa jedynie w dwóch wypadkach. Po pierwsze, gdy mamy do dyspozycji dane pomiarowe sieci rzeczywistej pracującej w warunkach analogicznych do zadanych w modelu symulacyjnym. Po drugie, gdy dysponujemy odpowiednim modelem analitycznym. W innych przypadkach możemy uważać otrzymane wyniki za lepsze lub gorsze przybliżenie symulowanego obiektu, którym w tym przypadku jest prosta (dwuwęzłowa) sieć lokalna.

Należy podkreślić, że modelowanie symulacyjne ma głównie charakter statystyczny. W trakcie eksperymentu symulacyjnego oblicza się zwykle momenty różnych zmiennych losowych. Poważnym utrudnieniem jest istnienie różnych powiązań statystycznych w samym modelu. Często też rozkłady obserwacji są niestacjonarne. Dodatkowo istnieje konieczność eliminacji wpływu warunków początkowych. Przyjęty sposób ominięcia stanu przejściowego omówiono w p. 4.

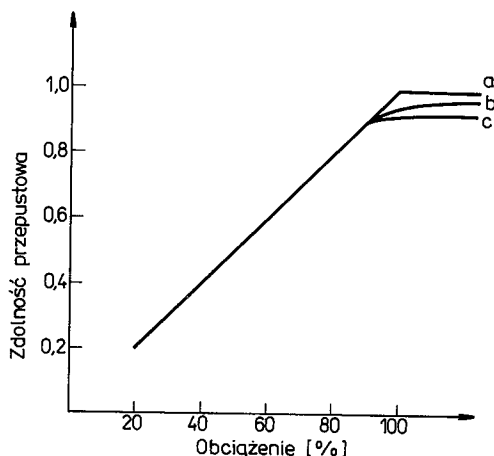
W praktyce spotyka się głównie dwie metody oceny dokładności wyników symulacji [3]. Pierwsza polega na wielokrotnym powtarzaniu przebiegu symulacyjnego. Jej zasadniczą wadą jest konieczność eliminowania wpływu warunków początkowych przy każdym powtórzeniu. Druga metoda, którą zastosowano w opracowanym programie, polega na pojedynczym, dostatecznie długim eksperymencie symulacyjnym. Przebieg dzielony jest na wystarczająco dużą liczbę części o równej długości. Oblicza się wartość średnią dla każdego odcinka, a poszczególne średnie odcinkowe traktuje się jako obserwacje niezależne. Estymatorem badanej zmiennej jest wartość średnia obliczona ze średnich odcinkowych. Przyjęcie założenia o niezależności i zastosowanie centralnego twierdzenia granicznego pozwala na przedstawienie obserwacji średnich odcinkowych za pomocą rozkładu normalnego. Umożliwia to zastosowanie znanych wzorów na estymator wariancji wartości średniej i obliczenie przedziału ufności. Przyjęcie założenia o niezależności jest usprawiedliwione, jeśli długość odcinka jest dostatecznie wielka. Występujące zjawisko autokorelacji polega

na istnieniu wpływu wartości danych dla jednego odcinka na wartość następnych danych (koniec jednego odcinka stanowi początek następnego). Wpływ ten maleje ze wzrostem odległości czasowej między danymi dla dwóch kolejnych odcinków. Po przekroczeniu pewnego przedziału czasowego można przyjąć, że może on być pominięty. Zgodnie z sugestiami w [3] przyjęto, że długość odcinka jest większa od przedziału między odcinkami. Ponadto długość tego odcinka jest znacznie większa od długości przedziału rozruchowego (początkowego).

## 6. EKSPERYMENTY SYMULACYJNE

Przedstawiony w p. 3 model symulacyjny służy do wymiany informacji między dwoma użytkownikami (np. mikrokomputerami) lokalnej sieci magistralowej. Stanowi on pierwszy krok w kierunku symulacji wielowzłowej sieci lokalnej. Zgodnie z przeprowadzonymi tam rozważaniami informacja na drodze od źródła do ujścia jest opóźniona na kilku etapach. Po pierwsze jest to związane z możliwym oczekiwaniem w komputerze źródłowym, czyli przed wejściem do właściwej podsieci komunikacyjnej. Następnie na trasie z komputera do adaptera. Z kolei w adapterze opóźnienie związane jest z tworzeniem ramki do wysłania oraz z dostępem do magistrali. W przypadku znacznego obciążenia magistrali to opóźnienie staje się istotne. Dalszymi składnikami zwiększającymi opóźnienie są: czas transmisji w magistrali, przetwarzanie ramki w adapterze odbierającym (początkowo założono, że wynosi zero) i ewentualnie oczekiwanie na możliwość przesłania w kanale adapter — komputer docelowy.

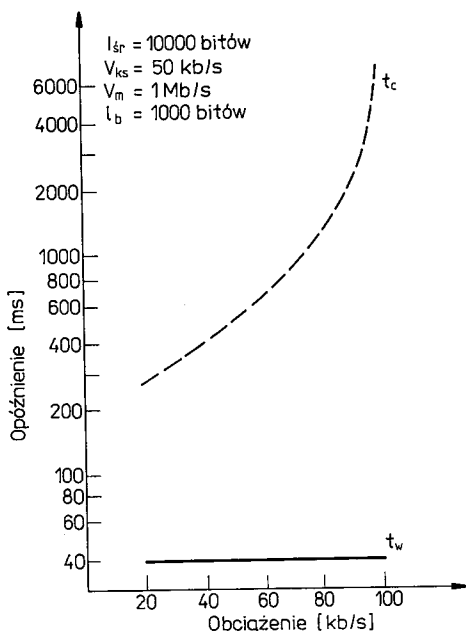
W pierwszym eksperymencie symulacyjnym badano charakterystykę zdolności przepustowej w funkcji obciążenia. Jest to bowiem zasadnicza zależność dla każdej sieci lokalnej [5]. W przypadku tworzenia modelu symulacyjnego można na tej podstawie ocenić w przybliżeniu jego zgodność z rzeczywistością. Uzyskane wyniki przedstawiono na rys. 4. Dla porównania pokazano tam także analogiczną charakte-



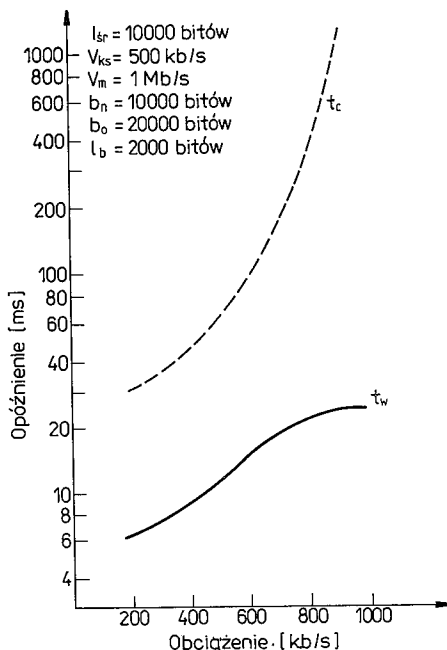
Rys. 4. Zdolność przepustowa w funkcji obciążenia; a) przypadek idealny, b) przebieg z pomiarów na podstawie [10], c) dla modelu symulacyjnego

rystykę otrzymaną na drodze pomiarów [10]. Na tej podstawie można stwierdzić, że podstawowe zależności dostępu rywalizacyjnego z wykrywaniem kolizji zostały odwzorowane prawidłowo w opracowanym modelu symulacyjnym.

Zwykle w sieciach lokalnych prędkości transmisji na styku komputer-adapter są znacznie mniejsze niż w magistrali. Fakt ten powoduje, że nawet przy znacznej liczbie stacji magistrala jest często niezbyt mocno obciążona. Potwierdzają to przeprowadzone eksperymenty symulacyjne. Do obliczeń tych przyjęto, że prędkość transmisji na styku komputer — adapter wynosi  $V_{ks} = 50$  kb/s, natomiast prędkość w magistrali  $v_m = 1$  Mb/s. Założono strumienie zgłoszeń o takiej samej intensywności  $\lambda_1 = \lambda_2$ , średnia długość wiadomości  $l_{sr} = 10\,000$  bitów, a długość bloku  $l_b = 1000$  bitów. Przeciętnie więc wiadomość składała się z dziesięciu bloków. Porównywano średni czas transmisji bloku (ramki)  $t_c$  ze średnim właściwym czasem transmisji bloku  $t_w$ . Wielkość  $t_c$  obejmuje okres czasu od momentu powstania (generacji) bloku do chwili zakończenia transmisji. Natomiast  $t_w$  liczony jest od rozpoczęcia wysłania bloku w komputerze źródłowym do zakończenia jego odbioru w komputerze docelowym. Otrzymane wyniki przedstawiono na rys. 5. Właściwy czas transmisji bloku ( $t_w$ ) jest praktycznie rzecz biorąc stały, gdyż magistrala jest bardzo słabo obciążona i blok po opuszczeniu komputera nie jest praktycznie opóźniany czekaniem w buforach na trasie swojej wędrówki. O opóźnieniu decyduje czas transmisji na poszczególnych odcinkach drogi. Natomiast średnie opóźnienie bloku ( $t_c$ ) jest dużo większe i szybko rośnie ze wzrostem obciążenia sieci. Decydującym czynnikiem jest oczekiwanie w kolejce przed wejściem do podsieci komunikacyjnej (bufor  $B_1$  na rys. 3). Nie występuje również opóźnienie związane z kolejkami w buforach adapterów.

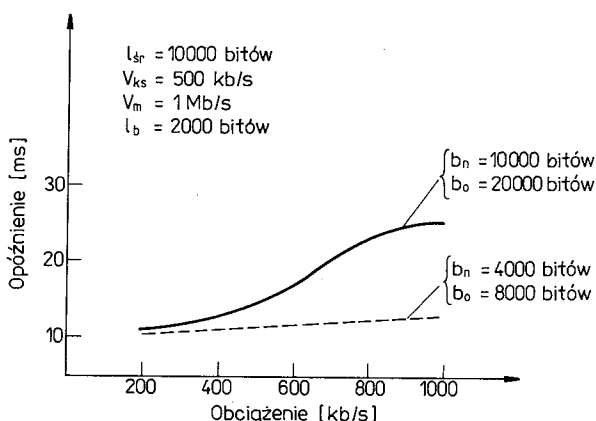


Rys. 5. Średnie opóźnienie bloków w funkcji obciążenia



Rys. 6. Średnie opóźnienie bloków w funkcji obciążenia

Sytuacja jest zupełnie inna, gdy maleje różnica między szybkością transmisji  $v_{ks}$  (na styku komputer-adapter) a  $v_m$  (w magistrali). Taka sytuacja może mieć również miejsce, gdy w sieci występują adaptory grupowe, do których podłączono kilka lub więcej urządzeń końcowych. Wówczas sumaryczny strumień zgłoszeń staje się porównywalny z prędkością w magistrali. Oznacza to możliwość powstawania kolejek w buforach nadawczych adapterów. Na rys. 6 przedstawiono wielkości  $t_c$  i  $t_w$  dla sieci, w której  $v_{ks} = 500$  kb/s, a  $v_m = 1$  Mb/s. Podano również wielkości buforów: nadawczego  $b_n = 10\,000$  bitów i odbiorczego  $b_o = 20\,000$  bitów. Jak widać, ze wzrostem strumienia zgłoszeń (obciążenie sieci) rośnie zarówno  $t_c$  jak i  $t_w$ . Wzrost  $t_w$  wynika z rosnącej zajętości magistrali, w związku z czym bloki (ramki) muszą coraz dłużej czekać w buforach adapterów. Można jednak zauważyć, że właściwy czas transmisji bloku stabilizuje się przy dużym obciążeniu. Oznacza to nasycenie przepustowości magistrali, a dalsze zwiększanie obciążenia powoduje tylko zwiększanie kolejek w źródle.



Rys. 7. Średnie opóźnienie właściwe bloku w funkcji obciążenia

Zbadano również — przykładowo — wpływ wielkości buforów w adapterach na opóźnienie przesyłania bloków. Oczywiście w sytuacji  $v_{ks} \ll v_m$  nie ma to praktycznie znaczenia, bo i tak w adapterze nie tworzą się kolejki. Takie porównanie ma sens, gdy prędkość transmisji na styku komputer — adapter jest niewiele mniejsza od prędkości w magistrali. Pokazano to na rys. 7. Zmniejszenie pojemności buforów powoduje, że wolniej rośnie średni właściwy czas transmisji bloku. Jest to związane z tym, że maleje opóźnienie wprowadzane, przez adapter, bo tworzą się w nim mniejsze kolejki. Oczywiście powoduje to pewne wydłużenie czekania wiadomości w źródle przy większym obciążeniu.

## 7. UWAGI KOŃCOWE

W pracy przedstawiono symulacyjne badanie wymiany informacji między dwoma stacjami w lokalnej sieci magistralowej z dostępem rywalizacyjnym typu CSMA-CD. Stanowi to pewien etap badań prowadzonych w Instytucie Podstaw Informatyki PAN

[8] nad możliwością zastosowania lokalnych sieci przetwarzania danych dla potrzeb medycznych systemów informatycznych. Przedstawiony model jest elastyczny i w swej prostej wersji umożliwia prowadzenie badań dla różnych konfiguracji — wariantowanie szybkości transmisji, pojemności buforów w adapterze czy modeli strumieni zgłoszeń.

Propozycja wykorzystania lokalnej sieci magistralowej do realizacji informatycznego systemu szpitalnego wychodzi na przeciw rosnącym wymaganiom w tej dziedzinie. Przyjęta koncepcja systemu o strukturze rozproszonej pozwala na lepszą realizację struktury informacyjnej szpitala zarówno w pionie diagnostyczno-lecznym, jak i w pionie administracyjno-ekonomicznym. System o strukturze rozproszonej, w odróżnieniu od systemu scentralizowanego, może funkcjonować, mimo awarii niektórych jego elementów, chociaż zwykle prowadzi to do częściowego ograniczenia jego funkcji użytkowych. Struktura rozproszona ma pod względem niezawodności wyraźną przewagę nad strukturą scentralizowaną choć jest trudniejsza w realizacji.

#### BIBLIOGRAFIA

1. L. Borzemski: *Developments in local area networks for data acquisition and process control*. International Conference on System Science, Wrocław, September 1986
2. L. Borzemski, J. Józefczyk, J. Świątek: *Wykorzystanie mikrokomputerów do analizy układów sterowania w pracach badawczych, projektowych oraz dydaktyce*: Materiały konferencji „Mikrokomputery w automatyce i technice systemów” cz. II, Wrocław 1984
3. G. Gordon: *System Simulation*. Prentice-Hall, Inc. 1969
4. A. Gościński, K. Zieliński: *Mikrokomputerowa sieć lokalna do sterowania*. Materiały konferencji „Mikrokomputery w automatyce i technice systemów” cz. III, Wrocław 1984
5. D. Hutchinson, J. Mariami, Shepherd (eds.): *Local Area Networks — An Advanced Course*. Lecture Notes in Computer Science, No 184, Springer Verlag, Heidelberg 1985
6. B. K. Jones: *A microcomputer network for metabolism studies*. Medical Informatics, 1985, Vol. 10, Nr 2, str. 199-200
7. I. Klimek-Grzesiak: *Wymagania stawiane lokalnym sieciom komputerowym przez informatyczny system szpitalny*. Prace Naukowe Centrum Obliczeniowego Politechniki Wrocławskiej, materiały konferencji „Sieci komputerowe, teoria, technika zastosowania”, Wrocław 1987
8. T. Nowicki, K. Antosik, T. Stachowiak: *Evaluation of the delay throughput characteristics in selected local teleprocessing networks*. System Science. 1988, Vol. 14, No 1, str. 63-71
9. H. J. Seelos: *Towards the morphology of medical information systems*. Medical Informatics, 1988, Vol. 13, No 2, str. 71-79
10. J. F. Shock, J. A. Hupp: *Measured Performance of an Ethernet Local Network*. Communications of the ACM, 1980, Vol. 23, No 12, str. 711-721
11. J. Świątek: *Identyfikacja dwustopniowa oraz jej zastosowania techniczne i biomedyczne*. Prace Naukowe Instytutu Sterowania i Techniki Systemów, Monografie Nr 2, Wrocław 1987
12. J. Villalabos, J. L. Manzano, M. Blazquez, J. Bolaños, S. Lubillo: *Computerized system in intensive care medicine*. Medical Informatics, 1986, Vol. 11, No 3, str. 269-275
13. W. B. Watson: *Simulation Study of the Traffic Dependent Performance of a Prioritized CSMA Broadcast Network*. Computer Networks, 1979, Vol. 3, No 3, str. 427-434
14. A. Wolisz, M. Piwowski: *Design Considerations and Logical Architecture of the LSK-NET Local Area Network*. Podstawy Sterowania, 1986, Tom 15, Zeszyt 1, str. 37-68
15. J. W. Yeh: *Simulation of Local Computer Networks — a Case Study*. Computer Networks 1979, Vol. 3, No 3, str. 401-417.

K. ANTOSIK, J. HOSZA, T. STACHOWIAK

SIMULATION OF DATA TRANSMISSION BETWEEN TWO NODS  
IN BUS TYPE LOCAL PROCESSING NETWORK

## Summary

The paper describes the simulation package for analysing data transmission problems between two nodes from bus type local area network with CSMA-CD access scheme. This is a stage of research work carried out, by the Institute of Computer Science of the Polish Academy of Sciences for the implementation of local area processing networks in the field of medicing.

The bus type local area network is presented, the simulation model is described and basic modelling problems are discussed. Next the method of verification of the results is presented. Finally, certain simulation experiments are described and their results commented.

# Szesnastostanowe modulatory mikrofalowe PSK oraz QAM

TADEUSZ MORAWSKI

*Instytut Radioelektroniki, Politechnika Warszawska*

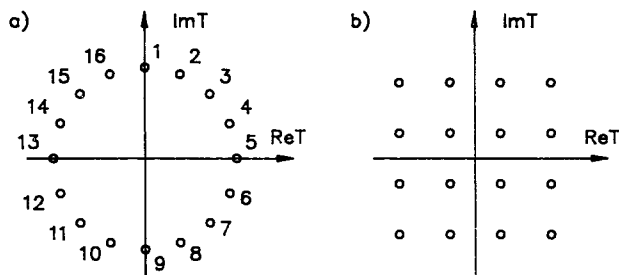
Otrzymano 1989.05.10

Autoryzowano do druku 1989.06.15

W pracy opisano metody analizy i projektowania szesnastostanowych modulatorów fazy PSK oraz modulatorów amplitudowo-fazowych QAM, stosowanych w systemach telekomunikacji mikrofalowej. Zanalizowano różne warianty modulatorów zawierające elementy pasywne w postaci sprzęgaczy 3 dB/90°, sumatorów oraz odcinków linii jednorodnych i sprzężonych oraz elementy kluczujące — diody PIN. Przedstawiono wyniki badań eksperymentalnych wybranych typów modulatorów PSK oraz QAM na pasma do 2,2 GHz.

## 1. WSTĘP

We współczesnych systemach telekomunikacyjnych (ziemskie i satelitarne linie radiowe) stosuje się bardzo często cyfrową modulację fazy PSK. Bardzo popularne są systemy z modulacją czterostanową [1]. Szybkość transmisji danych w systemie modulacji czterostanowej jest ograniczona przez czasy przełączeń elementów półprzewodnikowych (zwykle są to diody PIN, na etapie próbnym znajdują się modulatory z tranzystorami FET). Szybkość ta może ulec zwiększeniu bez postępu technologicznego w zakresie stosowanych elementów, jeżeli zwiększy się liczbę stanów fazowych modulatora fazy, np. do szesnastu stanów. W przypadku modulatorów szesnastostanowych skok fazy między kolejnymi stanami jest dość mały i wynosi  $22\frac{1}{2}^\circ$ . Utrzymanie w przybliżeniu stałego skoku fazy w pasmie pracy systemu (zwykle



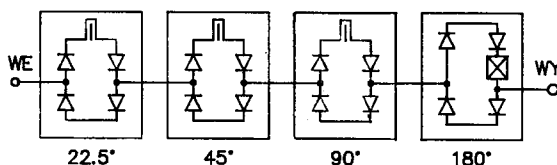
Rys. 1. Ilustracja systemów modulacji a) PSK; b) QAM

względna szerokość pasma jest rzędu  $\pm 10\%$  do  $\pm 20\%$ ) jest dość trudne, a ponadto w stosunku do modulacji czterostanowej pogarsza się rozróżnialność stanów w obecności szumów i zakłóceń. Dla zwiększenia tej odporności w najnowocześniejszych systemach proponuje się stosowanie modulacji amplitudowo-fazowej, tzw. quadrature amplitude modulation QAM, która pozwala na „rozsunięcia” punktów odpowiadających poszczególnym stanom modulatora na płaszczyźnie transmitancji zespolonej. Zilustrowano to na rys. 1. W artykule przedstawiono problemy projektowania szesnastostanowych modulatorów PSK i QAM oraz wyniki badań eksperymentalnych modeli modulatorów mikrofalowych na pasma częstotliwości  $0.5 \div 2$  GHz.

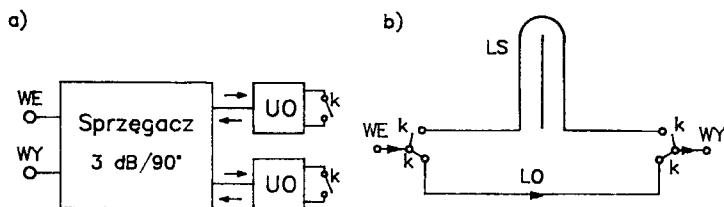
## 2. SZESNASTOSTANOWY MODULATOR FAZY

Wielostanowe modulatory fazy buduje się na ogół w ten sposób, że łączy się łańcuchowo obustronnie dopasowane modulatory dwustanowe, tak jak to pokazano na rys. 2.

Znanych jest wiele typów modulatorów dwustronnych [1, 2, 3, 4] przy czym najczęściej spotykane są modulatory odbiciowe ze sprzęgaczem  $3\text{ dB}/90^\circ$  lub modulatory transmisyjne typu „przełączana linia”. W tym ostatnim przypadku, dla uzyskania szerszego pasma pracy modulatora, jeden z odcinków linii jest tzw. przesuwnikiem Schiffmana [3], wykorzystującym zjawiska dyspersyjne w liniach sprzężonych.



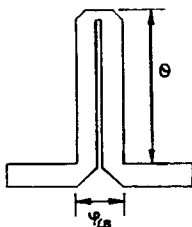
Rys. 2. Modulator 16-PSK złożony z czterech jednostek binarnych



Rys. 3. Typy modulatorów dwustanowych; a) odbiciowy ze sprzęgaczem  $3\text{ dB}/90^\circ$ , b) transmisyjny z przełączanymi liniami: linią sprzężoną LS oraz jednorodną linią odniesienia LO

Wymienione typy modulatorów przedstawiono na rys. 3. Jak widać z tego rysunku w modulatorze typu odbiciowego fala biegnie przez sprzęgacz oraz przez pasywne układy transformujące impedancję klucza dwukrotnie — od wejścia do kluczy i od kluczy do wyjścia, podczas gdy w układzie transmisyjnym fala biegnie bezpośrednio od

wejścia do wyjścia przez jedną z linii LS lub LO. Oznacza to, że w prawidłowo skonstruowanym dwustanowym modulatorze fazy typu odbiciowego, dopasowanym na wejściu i wyjściu, istnieją odcinki przewodnicy falowej, w których współczynnik fali stojącej jest bardzo duży. W przeciwieństwie do tego rozwiązania w modulatorze transmisyjnym wewnątrz modulatora (w odcinkach linii LS lub LO) współczynnik fali stojącej jest bliski jedności. Z powyższych powodów modulatory odbiciowe mogą wprowadzać większe straty niż modulatory transmisyjne. Modulatory te nie są więc preferowane do zastosowań w systemach wielostanowych, są one raczej stosowane do modulacji 4 PSK. Stosowane w modulacji wielostanowej modulatory transmisyjne są trudniejsze technologicznie, gdyż wymagane jest precyzyjne wykonanie wąskiej



Rys. 4. Sekcja linii sprężonych o długości elektrycznej  $\Theta$

szczeliny między liniami sprężonymi. W omawianym przypadku zaprojektowano modulator czterostanowy składający się z czterech jednostek binarnych połączonych łańcuchowo, tak jak to pokazano na rys. 2. Modulatory dwustanowe wykonano wykorzystując tzw. przesuwники Schiffmana (rys. 3b) z odcinkami połączonymi na końcu sekcji linii sprężonych w technice niesymetrycznych linii paskowych. Taką sekcję linii sprężonych pokazano na rys. 4.

Jeśli długość elektryczną linii jednorodnej oznaczono przez  $\Theta$ , to przesunięcie fazy  $\varphi_{LS}$  między wejściem i wyjściem sekcji linii sprężonych wyraża się wzorem (1):

$$\varphi_{LS} = \arccos \left( \frac{\varrho - \operatorname{tg}^2 \Theta}{\varrho + \operatorname{tg}^2 \Theta} \right), \quad (1)$$

gdzie

$$\varrho = \frac{Z_{oe}}{Z_{oo}} = \frac{1+k}{1-k} \quad (2)$$

$k$  — jest współczynnikiem sprzężenia między liniami,

$Z_{oe}$  — jest impedancją linii sprężonych przy pobudzeniu symetrycznym,

$Z_{oo}$  — jest impedancją przy pobudzeniu niesymetrycznym.

Charakterystyki fazowe sekcji linii sprężonych w odniesieniu do linii jednorodnej o impedancji charakterystycznej  $z = 1$  przedstawia rys. 5. Długość tej linii należy dobrać tak, aby uzyskać możliwie najmniejszy błąd fazy  $\delta\varphi$  w pasmie częstotliwości pracy układu,

$$\delta\varphi = |\Delta\varphi - \Delta\varphi_0| \quad (3)$$

gdzie:

$\Delta\varphi_0$  — projektowane przesunięcie fazy,

$\Delta\varphi = \varphi_{LS} - \varphi_{LO}$  — różnica faz linii sprężonej i odniesienia.

Na rys. 5 pokazano dwa sposoby dobierania długości linii odniesienia. Pierwszy,  $LO_1$  został zaproponowany przez Schiffmana [3], natomiast drugi,  $LO_2$  jest proponowanym pomysłem alternatywnym [5]. Długości linii odniesienia oraz odpowiadające im długości sprzężonych oblicza się z zależności:

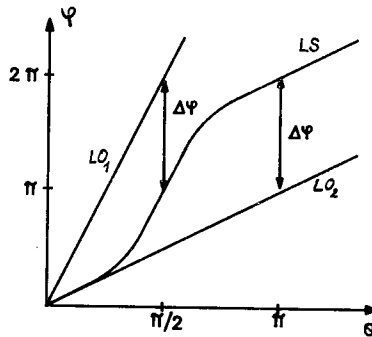
$$LO_1 = \frac{\left(1 + \frac{\Delta\varphi_0}{\pi}\right)c}{2f_0\sqrt{\varepsilon_{efo}}}, \quad (4)$$

$$LS_1 = \frac{c}{4f_0\sqrt{\varepsilon_{efs}}}, \quad (5)$$

$$LS_2 = \frac{c}{2f_0\sqrt{\varepsilon_{efs}}}, \quad (6)$$

$$LO_2 = \frac{\left(1 - \frac{\Delta\varphi_0}{2\pi}\right)c}{f_0\sqrt{\varepsilon_{efo}}}. \quad (7)$$

gdzie:  $\varepsilon_{ef}$  – efektywna stała elektryczna względna linii LO i LS.



Rys. 5. Charakterystyki fazowe sekcji linii sprzężonych

Analiza zależności (1) wykazuje, że istnieje pewne optymalne (ze względu na błąd fazy) sprzężenie linii. W przypadku maksymalnie płaskiej charakterystyki błęd fazy sprzężenie to charakteryzują następujące wartości współczynnika  $q$ :

$$q_{01} = \left(\frac{\Delta\varphi_0}{\pi} + 1\right)^2 \quad (8)$$

dla koncepcji Schiffmana, oraz:

$$q_{02} = \frac{2}{2 - \frac{\Delta\varphi_0}{\pi}}^2 \quad (9)$$

dla wersji proponowanej.

Z zależności (2, 8, 9) wynika, że osiąganie większych przesunięć fazy  $\Delta\varphi_0$  wymaga stosowania linii o większym współczynniku sprzężenia. W przypadku równomiernie falistej charakterystyki błędu fazy analiza zależności (1) jest bardziej skomplikowana, dlatego dokonuje się jej metodą numeryczną. Analiza ta wykazuje, że dla przesunięć fazy  $\Delta\varphi_0$  mniejszych od  $180^\circ$ , wartość optymalnego sprzężenia dla wersji  $LO_2$  jest mniejsza od klasycznej wersji  $LO_1$ :

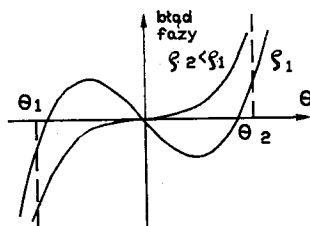
$$\varrho_{opt_2} < \varrho_{opt_1} \quad (10)$$

Jednocześnie błąd fazy w przypadku optymalnym jest nieco większy dla wersji  $LO_2$ :

$$\delta\varphi_{opt_2} > \delta\varphi_{opt_1} \quad (11)$$

Zalety przesuwnika z linią odniesienia  $LO_2$  widać w sytuacji, gdy ze względów technologicznych nie można uzyskać wystarczająco silnego sprzężenia. Błąd fazy w wersji z  $LO_2$  może być wtedy mniejszy od błędu w wersji z  $LO_1$ .

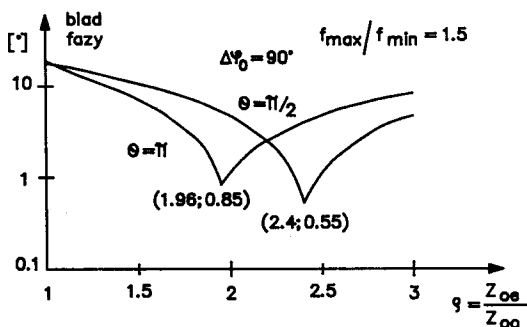
Drogą obliczeń numerycznych można się przekonać, że jeśli przekroczy się wartość współczynnika  $\varrho$  określoną równaniami (8), (9), to zależność błędu fazy od częstotliwości ma charakter krzywej typu „leżące S”, jak to pokazano na rys. 6, gdzie wartość  $\varrho_2$



Rys. 6. Błąd fazy modulatora z sekcją linii sprzężonych w funkcji częstotliwości

odpowiada charakterystyce maksymalnie płaskiej. Dla  $\varrho_1 > \varrho_2$  obserwuje się regularne zwiększenie się maksymalnego błędu fazy, przy jednoczesnym poszerzeniu się pasma częstotliwości.

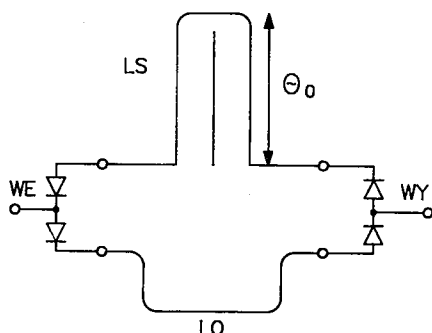
W omawianym przypadku wybrano do realizacji przesuwniki fazy pracujące w oktawowym pasmie pracy ( $f_{max}/f_{min} = 2$ ) i w tym ustalonym pasmie pracy minimalizowano błąd fazy poprzez dobór wartości  $\varrho$ . Dla zilustrowania zależności błędu fazy od



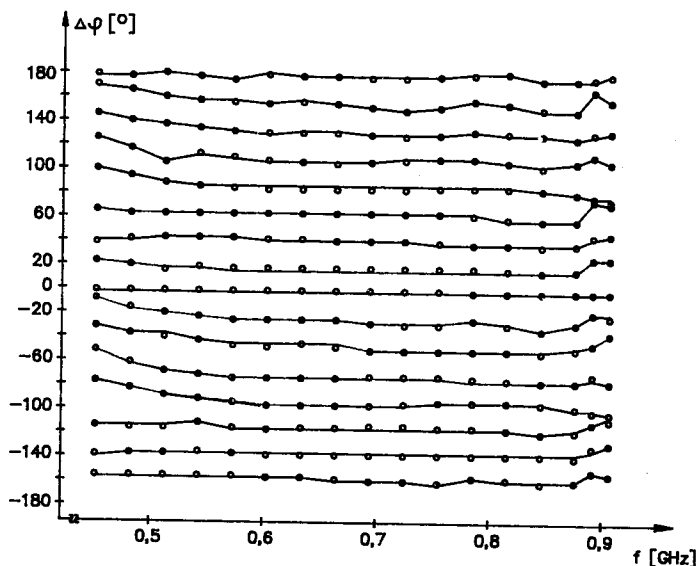
Rys. 7. Błąd fazy w funkcji współczynnika  $\varrho$

wartości  $\varrho$  dla przesuwnika o oktawowym pasmie pracy (tzn.  $\Theta_{min} = 60^\circ$ ,  $\Theta_{max} = 120^\circ$ ) wykonano odpowiednie obliczenia numeryczne. Wyniki zilustrowano na rys. 7.

Długości odcinków linii sprzężonej LS oraz linii odniesienia LO wyrażają się wzorami (4) i (6). Wykonano model modulatora szesnastostanowego w pasmie częstotliwości poniżej 1 GHz (rys. 2), [5]. Typowy modulator dwustanowy (jednostka z linią sprzężoną) bez układów polaryzacji diod pokazano na rys. 8. Modulatory dwustanowe 0/22.5°, 0/45° modulatora szesnastostanowego (rys. 1) wykonano w wersji klasycznej Schiffmana. Moduł 0/90°, gdzie występuje konieczność zastosowania silnie sprzężonych linii, wykonano w wersji z linią odniesienia typu LO<sub>2</sub>. W przypadku modułu 0/180° wymagany współczynnik sprzężenia okazał się zbyt duży dla obu wersji. Dlatego w module 0/180° wykorzystano skrzyżowanie przewodów („gorącego” z „masą”) w jednym z dwóch odcinków linii jednorodnych o jednakowej długości dla uzyskania odwrócenia fazy.



Rys. 8. Transmisyjny modulator dwustanowy



Rys. 9. Charakterystyki modulatora 16 PSK

Wyniki analizy przesuwników dwustanowych okazały się zgodne z wykonanymi eksperymentami, z tym, że przy częstotliwościach zbliżonych do  $f_{max}$  obserwuje się wyraźną odchyłkę fazy, spowodowaną rezonansem odcinków linii oraz pojemnością pasożytniczą diod PIN (przy projektowaniu zakładano, że diody są idealnymi kluczami).

Charakterystyki fazowe całego modulatora szesnastostanowego pokazano na rys. 9. Jak widać z rys. 9 osiągnięto mały błąd fazy całego modulatora. W pasmie  $0.5 \div 0.85$  GHz błąd ten nie przekracza  $3^\circ$ .

Z pomiarów amplitudowych wynika, że modulator szesnastostanowy jest dobrze zrównoważony — szkodliwa modulacja amplitudy nie przekracza 6%. Straty modulatorów dwustanowych z liniami sprzężonymi zgodnie z przewidywaniami są małe, pomimo stosowania najtańszych i dość stratnych elementów (laminatu i diod). Wynoszą one około 2 dB na jednostkę binarną. Całkowite straty modulatora szesnastostanowego rosną z częstotliwością. Dla  $f \leq 0.7$  GHz nie przekraczają one 8 dB.

Podkreślić należy, że osiągnięte prawie oktawowo pasmo pracy modulatora jest bardzo szerokie. W wielu zastosowaniach telekomunikacyjnych nie jest potrzebne aż tak szerokie pasmo (np. w radioliniach może wystarczyć pasmo o względnej szerokości  $\pm 5\%$ ). Zaprojektowanie modulatora na węższe pasmo częstotliwości umożliwia zmniejszenie błędów fazowych, a także zmniejszenie wartości współczynników sprzężenia linii, co ułatwia realizację układu.

### 3. SZESNASTOSTANOWY MODULATOR QAM

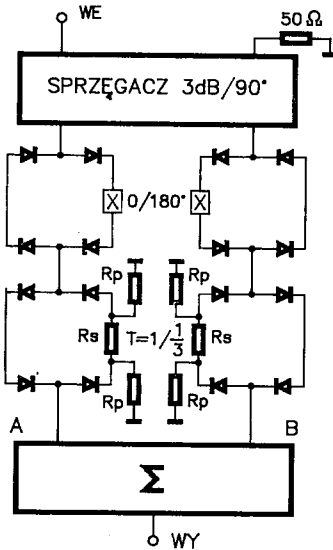
Szesnastostanowy modulator QAM można zbudować wykorzystując sprzęgacz kwadraturowy 3 dB oraz odpowiednie przełączniki fazy ( $0/180^\circ$ ) i amplitudy (0 dB/—9.54 dB) oraz sumując tak przełączane sygnały w sumatorze, jak to pokazano na rys. 10. Jako przełącznik fazy można zastosować modulator binarny typu transmisyjnego ze skrzyżowanymi przewodami (sekcja  $0/180^\circ$  z rys. 2). Innym rozwiązaniem jest przełącznik  $0/180^\circ$  w postaci sprzęgacza kwadraturowego, do którego dwóch gałęzi dołączono diody PIN. Przełączanie takich diod zapewnia  $180^\circ$  skok fazy, jeśli tylko ich parametry pasożytnicze L, C spełniają zależność:

$$\sqrt{L/C} = Z_0 = 50 \, \Omega, \quad (12)$$

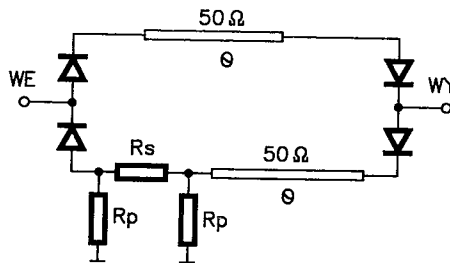
gdzie  $Z_0 = 50 \, \Omega$  jest impedancją odniesienia.

Przełączniki amplitudy są układami złożonymi z rezystorów i diod PIN. Bardzo krytyczne są błędy fazowe przełączników amplitudy, wprowadzane przez pasożytnicze elementy L, C diod PIN. Analiza różnych rozwiązań przełączników prowadzi do wniosku, że dobrymi parametrami charakteryzują się przełączniki transmisyjne, w których błędy fazowe kompensuje się przez umieszczenie w obu gałęziach przełącznika odcinków linii o takich samych, odpowiednio dobranych długościach — rys. 11.

Długość elektryczną kompensujących odcinków linii  $\Theta$  dobiera się tak, aby macierz admitancyjna tej gałęzi przełącznika, która zawiera diody PIN w stanie



Rys. 10. Modulator 16 QAM z przełącznikami fazy  $0/180^\circ$  oraz amplitudy  $\left(T = 1, \frac{1}{3}\right)$



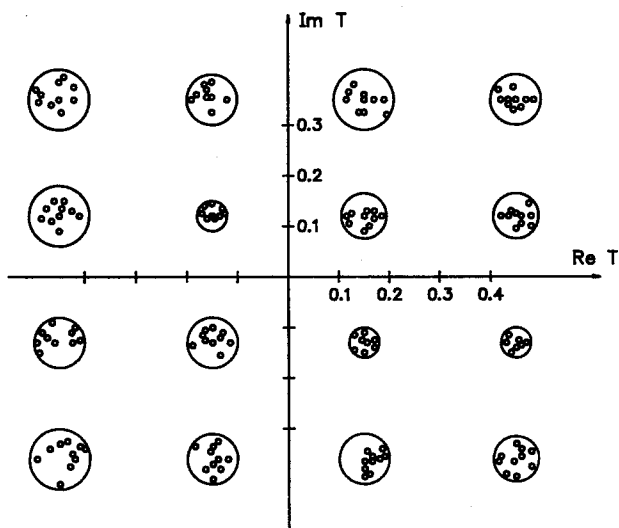
Rys. 11. Przełącznik amplitudy z odcinkami linii o długości elektrycznej  $\Theta$  minimalizującymi błędy fazowe

zaporowym (diody te są wówczas scharakteryzowane przez reaktancję pojemnościową  $X_c$ ) miała wyrazi jak najmniejsze i przez to jak najmniejszy wpływ na transmisję sygnału, która odbywa się przez gałąź zawierającą diody w stanie przewodzenia. Wartość  $\Theta$  oblicza się ze wzoru 6:

$$\Theta = \arctg \frac{|X_c|^2 - 1}{2|X_c|}. \quad (13)$$

Skonstruowano dwa modele szesnastostanowych modulatorów QAM na pasmo do 1 GHz [7]. W modelu o szerokim pasmie pracy zastosowano międzypalczasty sprzęgacz Langa. Uzyskano dobre wyniki w bardzo szerokim dwuoktawowym pasmie pracy  $0.19 \div 0.75$  GHz. Wyniki eksperymentalnego badania modulatora pokazano na rys. 12, na którym, dla uniknięcia obracania się rysunku w szerokim pasmie częstotliwości umieszczono transmitancje zredukowane w stosunku do średniej geometrycznej czterech transmitancji o największych modułach. W wielu zastosowaniach telekomunikacyjnych wystarczające jest znacznie węższe pasmo pracy, które można

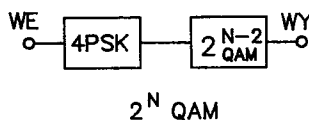
uzyskać stosując prostszy i łatwiejszy technologicznie sprzęgacz dwu- lub trójgałęziowy. Wykonano model modulatora ze sprzęgaczem trójgałęziowym na pasmo  $0.55 \div 0.8$  GHz. Osiągnięto straty nie większe niż  $6 \text{ dB} \pm 1 \text{ dB}$  oraz błędy fazy nie większe niż  $\pm 6^\circ$ . Skonstruowano też próbny model modulatora QAM na typowe pasmo radiolinii  $1.8 \div 2.2$  GHz. W modulatorze tym, oprócz sprzęgacza trójgałęziowego na wejściu, wykorzystano przełączniki  $0/180^\circ$  w postaci sprzęgaczy trójgałęziowych obciążonych



Rys. 12. Wyniki badań modulatora 16 QAM

dwiema diodami PIN. Osiągnięto również pozytywne wyniki eksperymentu, straty były nie większe niż  $6 \text{ dB} \pm 1 \text{ dB}$ , a błędy fazowe  $6^\circ$ . Należy podkreślić, że ostatni z układów modulatora 16 QAM zawiera więcej sprzęgaczy (trzy) niż układ opisany poprzednio (jeden sprzęgacz) ale zamiast 16 diod PIN wymaga mniejszej liczby diod — 12.

Innym rozwiązaniem układu modulatora 16 QAM jest łańcuchowe połączenie modulatora fazy 4 PSK oraz modulatora amplitudowo-fazowego 4 QAM — rys. 13.



Rys. 13. Modulator  $2^N$  QAM złożony z modulatorów 4 PSK oraz  $2^{N-2}$  QAM

Do zbudowania tego modulatora wykorzystano skonstruowane tak jak w poprzednio opisanych modelach sprzęgacze trójgałęziowe (SK1/SK2/SK3), przełączniki amplitudy oraz sumator. Układy odbijające U01/U02, zawierające diody PIN i odpowiednie transformatory impedancji dla uzyskania skoków fazy  $0/90^\circ$  oraz  $0/180^\circ$ , zaprojektowano w/g metody opisanej w [4], uwzględniającej teoretycznie obliczone

parametry zaprojektowanego sprzęgacza oraz zmierzone parametry diod PIN. Uzyskano małe błędy modulatora 4 PSK, [1], oraz dobre wyniki dla całego modulatora 16 QAM, porównywalne z przypadkiem opisanym poprzednio.

## PODSUMOWANIE

W pracy opisano zagadnienia, związane z projektowaniem modulatorów 16-PSK oraz 16-QAM, przy czym elementami tych złożonych układów były w różnych wariantach zarówno modulatory binarne transmisyjne jak i odbiciowe, przełączniki fazy i amplitudy. W przypadku transmisyjnego modulatora fazy najtrudniejsze problemy są związane z koniecznością uzyskania bardzo wąskiej szczeliny w sekcji linii sprzężonych. W przypadku modulatorów QAM krytyczne są błędy fazowe wprowadzone przez przełączniki amplitudy, które jednak zostały zminimalizowane przez zastosowanie kompensujących odcinków linii 50Ω. Wykonane badania eksperymentalne wykazały, że w obu przypadkach modulatorów PSK oraz QAM możliwe jest osiągnięcie dobrych parametrów w szerokich pasmach częstotliwości.

## BIBLIOGRAFIA

1. T. Morawski, J. Zborowska: *Mikrofalowe odbiciowe czterostanowe modulatory fazy*. Przegląd Telekomunikacyjny, 1987, nr 8, ss. 217—235
2. T. Morawski, K. Modelski: *Mikrofalowe modulatory i przesuwniki fazy z diodami półprzewodnikowymi*. PWN, Warszawa—Wrocław, 1984
3. B. M. Schiffman: *A New Class of Broad-Band Microwave 90-degree Phase Shifters*, IRE, Trans. on MTT, No 4, 1958
4. J. Komisarczuk: *Broad-band phase modulator with a branch line coupler-transformers*. Proc. of 9-th European Microwave Conference, Brighton, 1979
5. T. Morawski, J. Zborowska, C. Mroczkowski, L. Skupiński: *Model 16-stanowego modulatora fazy PSK* — Krajowa Konferencja Mikrofalowa MIKON, Gdańsk, 1988, ss. 20—25, tom I
6. J. Zborowska: *Przełączniki amplitudy do modulatorów QAM*, zgłoszenie na Krajowe Sympozjum Telekomunikacji, Bydgoszcz 1989
7. T. Morawski, J. Zborowska, C. Mroczkowski, J. Stypułkowski: *Wielostanowe modulatory amplitudowo-fazowe na pasmo VHF*. Krajowe Sympozjum Telekomunikacji, Bydgoszcz 1988, ss. 80—85, tom D

T. MORAWSKI

## SIXTEEN-STATE MICROWAVE MODULATORS PSK AND QAM

### Summary

Methods of analysing and designing of sixteen-state phase modulators PSK and amplitude-phase modulators QAM, applied in microwave telecommunication systems, are described. Different modulator variants, containing passive elements in the form of 3dB/90° couplers, adders as well as homogeneous line segments and keying elements — diodes PIN, are analysed. Results of experimental research on selected types of modulators PSK and QAM for bands up to 2.2 MHz are presented.

# Zakłócenia przemysłowe w miastach polskich w zakresie fal średnich radiofonicznych

ZBIGNIEW RYMAROWICZ

*Instytut Łączności, Wrocław*

*Otrzymano 1988.10.14*

*Autoryzowano do druku 1989.05.05*

W pracy omówiono badania zakłóceń przemysłowych na terenach mieszkaniowych miast. Objęto nimi 15% miast Polski. Badano zależność poziomu zakłóceń od czasu, częstotliwości, rodzaju zabudowy, wielkości miasta i stopnia jego uprzemysłowienia. Przedstawiono model teoretyczny i wyniki badań eksperymentalnych.

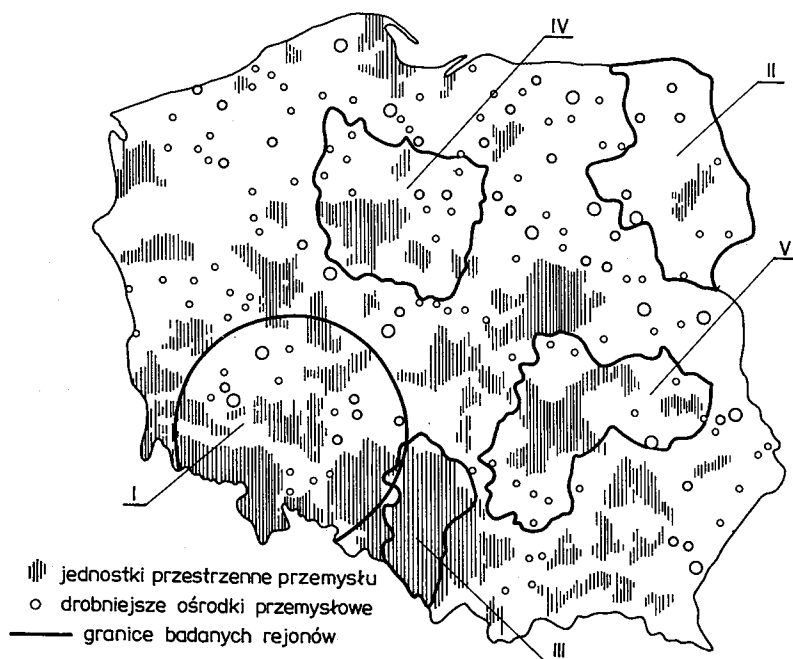
## 1. WSTĘP

Zakłócenia przemysłowe są jednym z zasadniczych czynników ograniczających zasięgi służb radiodfuzyjnych w skupiskach miejskich. Sprawa ta znana była od dawna, jednak zaczęła nabierać dużego znaczenia w miarę postępującego rozwoju miast i wzrastającej liczby urządzeń elektrycznych — źródeł zakłóceń. Spowodowało to, że organizacje międzynarodowe zajmujące się problemami związanymi z gospodarką widmem elektromagnetycznym zagadnienie to uznały za ważne i włączyły do swoich programów badań [1].

Dla oceny poziomu zakłóceń w obszarach umiastowionych i w przestrzeni ponad nimi prowadzone są od wielu lat, w niektórych krajach, badania pokrywające bardzo szeroki zakres częstotliwości [2—11]. Większość z tych prac dotyczy częstotliwości powyżej 20 MHz. Ich autorzy badali głównie zależności między poziomem zakłóceń a częstotliwością lub stopniem urbanizacji rozpatrywanego obszaru. Zawarte w tych pracach wyniki różnią się znacznie między sobą. Przyczyny tych rozbieżności należy szukać przede wszystkim w różnicach w rozwoju technicznym krajów, w których prowadzono badania, a także zastosowanej metodyce i aparaturze pomiarowej.

Badania omówione w niniejszej pracy zapoczątkowano w 1969 roku pomiarami we Wrocławiu. Kontynuowano je następnie w latach siedemdziesiątych w miastach na terenie 5 rejonów o różnym stopniu uprzemysłowienia (rys. 1). Miały one dostarczyć informacji o poziomach zakłóceń występujących na terenach mieszkaniowych miast oraz ustalić jego związek z rodzajem zabudowy, wielkością miasta i stopniem uprzemysłowienia.

Niektóre fragmenty tych badań były prezentowane sukcesywnie w poprzednich



Rys. 1. Badane rejonu na tle rozmieszczenia przemysłu w kraju

latach [12—19], lecz po zebraniu całego materiału doświadczalnego wyniki pomiarów przeliczono ponownie, wprowadzając jednolity opis procesu zakłóceń za pomocą parametrów rozkładu prawdopodobieństwa przekroczeń. Wyniki te przedstawiono w niniejszej pracy.

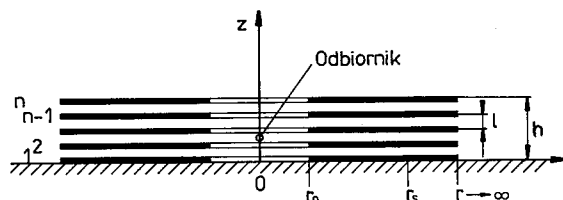
W pracy rozpatrzono również model miasta traktowany jako skupisko wielu źródeł zakłóceń. Przyjęte w modelu zależności związane ze strukturą zabudowy są bardzo uproszczone. Mimo to przeprowadzone rozważania umożliwiają lepsze zrozumienie zachodzących zjawisk i są pomocne przy interpretacji wyników badań eksperymentalnych. W prezentowanym modelu jak i badaniach rodzaj zabudowy określono za pomocą wskaźnika gęstości zaludnienia terenów mieszkaniowych netto<sup>1)</sup>. Wskaźnik ten charakteryzujący pewien obszar miasta uznano za przybliżoną miarę liczby źródeł zakłóceń, jak i przeszkód na drodze propagacji zakłóceń.

## 2. MODEL

Na odbiornik w mieście oddziałują jednocześnie wiele różnego rodzaju urządzeń, które mogą promieniować energię elektromagnetyczną w sposób niezamierzony zarówno w szerokim zakresie częstotliwości jak i dyskretne linie widmowe. Mechanizm powstawania w nich zakłóceń i przenikania ich do otoczenia może być różny i

<sup>1)</sup> Wskaźnik gęstości zaludnienia jest to liczba mieszkańców przypadająca na 1 ha terenu mieszkaniowego. Wskaźnik ten może być obliczany brutto, tzn. dla całego terenu z ulicami, placami, zieleńcami itp. albo obliczany netto tzn. dla czystych terenów mieszkaniowych [20].

szczegółowa ich analiza w rozpatrywanym modelu nie jest konieczna. W naszych rozważaniach wystarczy że przyjmimy iż urządzenia te, nazywane źródłami zakłóceń, w jednostce czasu emitują energię  $w_0$ . Ilość tej energii może być różna przy różnych częstotliwościach. Założymy następnie, że źródła te rozmieszczone są z jednakową gęstością powierzchniową  $\delta_s$  na  $n$  poziomach (kondygnacjach budynku). Odbiornik umieścimy pośród źródeł zakłóceń jak pokazano na rys. 2.



Rys. 2. Model miasta

Zakłócenia w pasmie poniżej 30 MHz zasadniczo rozchodzą się wzdłuż przewodów sieci elektrycznej lub innych struktur, wykonanych z przewodników prądu, z nią sprzężonych. Następnie przenikają one do odbiornika przez układy zasilania lub na skutek bezpośredniego sprzężenia anteny odbiorczej z siecią. Biorąc to pod uwagę zakładamy, że strumień energii zakłóceń dopływające do odbiornika od pojedynczych źródeł mają gęstość

$$S = \frac{w_0}{2\pi rh} \exp(-\alpha r) \quad (1)$$

gdzie:

$w_0$  – energia wydzielona przez źródło w jednostce czasu,

$r$  – odległość źródła od odbiornika,

$h$  – wysokość zabudowy,

$\alpha$  – współczynnik tłumienia.

Między współczynnikiem tłumienia  $\alpha$  i rodzajem zabudowy wyrażonym za pomocą gęstości zaludnienia terenów mieszkaniowych netto  $\varrho$  zachodzi następujący związek [21, 22]

$$\alpha = c_1 \varrho^{2/3} \text{ [dB/m]}, \quad (2)$$

gdzie  $c_1$  jest stałą.

Jeśli przyjmimy, że energia absorbowana przez odbiornik ( $w$ ) jest proporcjonalna do gęstości strumienia ( $S$ )

$$w = k S \quad (3)$$

to wówczas wypadkowa energia od wszystkich źródeł zakłóceń wyniesie

$$w_s = k \frac{w_0 n \sigma_s}{2\pi h} \int_{r_0}^{\infty} \int_0^{2\pi} \exp(-\alpha r) d\varphi dr, \quad (4)$$

skąd po wykonaniu całkowania mamy

$$w_s = k \frac{w_0 n \sigma_s}{\alpha h} \exp(-\alpha r_0), \quad (5)$$

gdzie  $r_0$  jest promieniem enklawy, która nie zawiera w swym wnętrzu źródeł zakłóceń. Promień tej enklawy wyznaczmy przyjmując, iż powierzchnia jej jest odwrotnie proporcjonalna do gęstości źródeł zakłóceń  $\sigma_s$ , skąd  $r_0 = c_2 \sigma_s^{-1/2}$  gdzie  $c_2$  jest stałą.

Jeśli oznaczymy przez  $V = n \sigma_s / \varrho$  liczbę źródeł zakłóceń przypadającą na mieszkańca oraz dla uproszczenia założymy, że wysokość zabudowy  $h = c_3 \varrho^{1/3}$  gdzie  $c_3$  jest stałą, to wówczas równanie (5) przyjmie postać

$$w_s = k_1 w_0 \Theta, \quad (6a)$$

w której

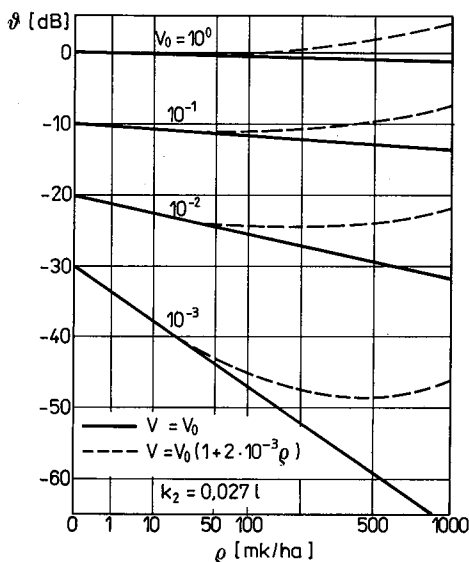
$$\Theta = V \exp(-k_2 V^{-1/2} \varrho^{1/3}), \quad (6b)$$

gdzie:

$$k_1 = k / (c_1 c_3),$$

$$k_2 = c_1 c_2 (c_3 / l)^{1/2}, \text{ przy czym } l \text{ jest wysokością jednej kondygnacji budynku.}$$

Na rys. 3 pokazano zależność  $\vartheta = 10 \lg \Theta$  od gęstości zaludnienia terenów mieszkaniowych  $\varrho$  dla przykładowo wybranych wartości  $V$ . Z rysunku tego wynika, że poziom zakłóceń będzie maleć ze wzrostem gęstości zaludnienia przy stałej liczbie źródeł zakłóceń przypadającej na mieszkańca. Wzrastać zaś będzie kiedy liczba źródeł  $V$  wzrasta z gęstością zaludnienia, a wpływ tego wzrostu na wypadkową energię zakłóceń  $w_s$  przewyższy wpływ tłumienia wnoszonego przez zabudowę.



Rys. 3. Zależność  $\vartheta$  od gęstości zaludnienia terenów mieszkaniowych

Strefę spływu zakłóceń wyznaczmy dzieląc cały obszar emitujący energię zakłóceń na dwie części. Część pierwsza rozciągać się będzie od  $r_0$  do  $r_s$ , a część druga od  $r_s$  aż do nieskończoności. Niech energia zakłóceń  $w_{s1}$  odebrana ze strefy pierwszej z błędem

$$\frac{w_s - w_{s1}}{w_s} \leq \varepsilon, \quad 0 \leq \varepsilon \leq 1 \quad (7)$$

będzie równa energii zakłóceń  $w_s$  odebranej od wszystkich źródeł zakłóceń z całego obszaru nieograniczonego. Dalej część pierwszą nazywać będziemy strefą spływu zakłóceń a jej promień  $r_s$  promieniem spływu zakłóceń. Uwzględniając rów. (4) uzyskuje się

$$\varepsilon \geq \exp[-\alpha(r_s - r_0)],$$

skąd promień strefy spływu zakłóceń wyniesie

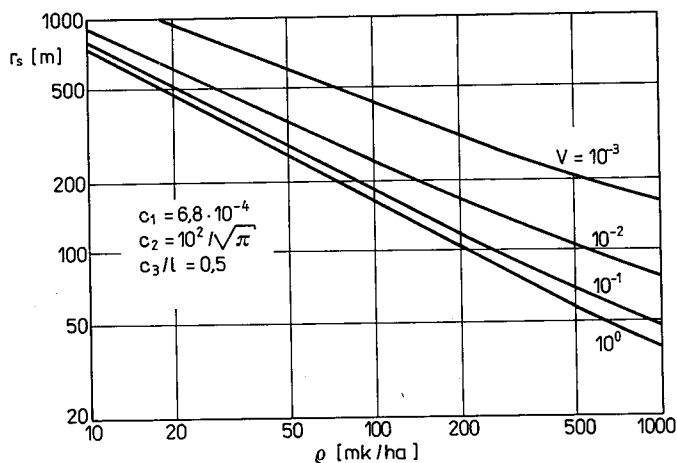
$$r_s = r_0 - \frac{\ln \varepsilon}{\alpha}, \quad (8)$$

a po podstawieniu wyrażen na  $\alpha$  i  $r_0$  ostatecznie otrzymujemy

$$r_s = k_3 V^{-1/2} Q^{-1/3} - k_4 (\ln \varepsilon) Q^{-2/3} \quad (9)$$

gdzie  $k_3 = c_2(c_3/l)^{1/2}$ ,  $k_4 = 1/c_1$

Przebieg powyższej zależności dla przykładowo wybranych wartości  $V$  został przedstawiony na rys. 4. Przy wykreślaniu krzywych przyjęto  $\varepsilon = 0,1$  co odpowiada wartości około  $-0,5$  dB. Z rys. 4 wynika, że promień strefy spływu zakłóceń przy poczynionych założeniach wynosi w przybliżeniu od 80 do 500 metrów, przy czym wartość mniejsza dotyczy zabudowy zwartej o dużej gęstości zaludnienia, a wartość większa zabudowy luźnej o gęstości około 30 mk/ha. Wyniki te wskazują, iż na



Rys. 4. Promień strefy spływu zakłóceń w funkcji gęstości zaludnienia terenów mieszkaniowych

wielkość odbieranej energii zakłóceń największy wpływ wywiera najbliższe otoczenie rozpatrywanego punktu.

Równanie (5), określające wielkość wypadkowej energii zakłóceń, odnosi się do obszarów o jednolitym rodzaju zabudowy dla których gęstość zaludnienia można uważać za stałą. Biorąc pod uwagę całe miasto, gdzie można wyróżnić szereg rejonów lub stref zabudowy różniących się od innych rodzajem zabudowy, za miarę wypadkowej energii zakłóceń przyjmujemy wartość średnią ważoną

$$\bar{w}_s = \sum_i \eta_i w_{si} \quad (10)$$

gdzie  $w_{si}$  jest energią zakłóceń wyrażoną w dB,  $\eta_i = m_i/m$  jest wagą wyrażającą udział ludności  $i$ -tego rejonu zabudowy w całości miasta,  $m_i$  jest jego liczbą ludności, a  $m$  liczbą ludności miasta. Uwzględniając wyrażenie (5) w (10) otrzymujemy

$$\bar{w}_s = 10 \lg k_1 w_0 + \bar{\mathcal{G}} \quad (11)$$

gdzie

$$\bar{\mathcal{G}} = 10 \sum_i \eta_i \lg V_i - 4.34 k_2 \sum_i \eta_i V_i^{-1/2} q_i^{1/3}. \quad (12)$$

Jeśli przyjmujemy  $V_1 = V_2 = \dots = V$ , to wtedy

$$\bar{\mathcal{G}} = 10 \lg V - 4.34 k_2 V^{-1/2} \cdot \overline{q^{1/3}} \quad (13)$$

gdzie podstawiono

$$\overline{q^{1/3}} = \sum_i \eta_i q_i^{1/3}.$$

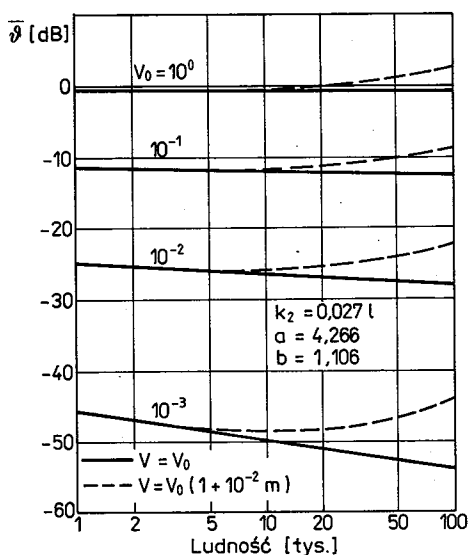
Struktura zabudowy przestrzennej miasta zmienia się ze zmianą jego wielkości. Zmiany te wpływają na zmianę  $\overline{q^{1/3}}$ , co z kolei ma wpływ na wartość  $\bar{w}_s$ . Opierając się na danych urbanistycznych dla miast polskich, zależność  $\overline{q^{1/3}}$  od wielkości miasta opisano wzorem

$$\overline{q^{1/3}} = a + b \lg m \quad (14)$$

gdzie  $a$  i  $b$  są stałymi,  $m$  jest liczbą ludności miasta w tys. Podstawiając powyższe wyrażenie do równ. (13) otrzymujemy

$$\bar{\mathcal{G}} = 10 \lg V - 4.34 k_2 V^{-1/2} (a + b \lg m). \quad (15)$$

Przebieg tej zależności dla przykładowo wybranych wartości  $V$  pokazano na rys. 5. Z rysunku tego widać, iż przy stałej liczbie źródeł zakłóceń przypadającej na mieszkańca

Rys. 5. Zależność  $\bar{\bar{J}}$  od wielkości miasta

średnia ważona energia zakłóceń  $\bar{\bar{w}}_s$  będzie maleć wraz z wielkością miasta. Wzrastać ona będzie kiedy liczba źródeł zakłóceń  $V$  równocześnie będzie wzrastała z wielkością miasta, a wpływ tego wzrostu na  $\bar{\bar{w}}_s$  przewyższy wpływ tłumienia wnoszonego przez zabudowę.

### 3. BADANIA EKSPERYMENTALNE

#### 3.1. OPIS POMIARÓW

Pomiary we Wrocławiu miały dostarczyć informacji o poziomie zakłóceń w dużym mieście i ustalić związek jego wartości przede wszystkim z charakterem i rodzajem zabudowy. Objęto nimi 35 rejonów miasta różniących się strukturą przestrzenną, intensywnością zabudowy i gęstością zaludnienia terenów mieszkaniowych.

Pomiary w pozostałych miastach miały na celu uchwycenie wpływu wielkości miasta i stopnia jego uprzemysłowienia na poziom zakłóceń, jak również zorientowanie się o wartościach jaki on przybiera. Do badań wybrano 123 miasta o liczbie ludności od 1 do 600 tysięcy. Większość stanowiły miasta małe i średnie, jakie często znajdują się na krańcach lokalnych stacji radiofonicznych, gdzie wpływ zakłóceń na odbiór jest bardzo odczuwalny. Miasta objęte badaniami stanowiły 15% wszystkich miast Polski, co można uznać za wystarczająco reprezentatywną próbkę dla całej populacji miast naszego kraju.

Miejsca pomiaru usytuowane były na ulicach lub placach terenów mieszkaniowych. Przy ich doborze zwracano uwagę, aby w pobliżu nie występowało miejscowe, dominujące źródło zakłóceń, które mogłoby maskować zakłócenia pochodzące od miasta jako całości. W wybranych miejscach, obserwacje prowadzono we wszystkie dni tygodnia z wyjątkiem sobót i niedziel, w godzinach pomiędzy 16 a 20.

Mierzono składową pionową natężenia pola elektrycznego na częstotliwościach 0,5, 1,35 i 1,8 MHz we Wrocławiu, a 0,5 i 1,6 MHz w pozostałych miastach. Do pomiaru używano anteny prętowej pionowej o długości 2,5 metra. Podstawa anteny znajdowała się na wysokości 1,5 metra nad ziemią. Cechowanie jej przeprowadzono w jednorodnym polu fali płaskiej.

W pomiarach stosowano miernik zakłóceń oparty na zaleceniach CISPR z dedektorem wartości quasi-szczytowej. Rozróżniano dwie wielkości: zakłócenia ciągle tworzące tło i krótkotrwałe impulsy (pojedyncze lub w postaci paczek) o poziomie znacznie wyższym od tła <sup>2)</sup>. W taki sposób zebrano cały obszerny materiał eksperymentalny (około 10 tysięcy wyników).

Wyniki uzyskane z tych pomiarów poddano analizie regresji wielorakiej, badając wpływ branych pod uwagę czynników na poziom zakłóceń. Estymację współczynników równania przeprowadzono metodą najmniejszych kwadratów.

Pomiary w godzinach 16—20 nie dają pełnego obrazu niezbędnego do oceny zakłóceń. Stąd wynikła potrzeba uzupełnienia tych badań obserwacjami pełnodobowymi. Badania te przeprowadzono w dwu charakterystycznych rejonach kraju, wyraźnie różniących się stopniem uprzemysłowienia i strukturą ludności. Do badań tych wybrano rejon II zaliczany do jednego z najslabiej uprzemysłowionych i rejon III szczególnie wysoko uprzemysłowiony, które pokazano na tle rozmieszczenia przemysłu w kraju na rys. 1. W rejonie II pomiarami objęto 7 miast o wielkości od 1 do 50 tysięcy mieszkańców, a w rejonie III 11 miast o wielkości od 1 do 200 tysięcy mieszkańców. Obserwacje prowadzono na częstotliwości 0,5 MHz w dni powszednie w ciągu całej doby, natomiast w niedziele w godzinach między 9 i 12 oraz 15 i 18.

Mierzono czas i ilość przekroczeń przez wartość quasi-szczytową nastawionego poziomu, jednocześnie wykonując pomiar tła i krótkotrwałych impulsów zakłóceń. Pomiar ten wykonywano za pomocą miernika zakłóceń i dołączonego do niego analizatora trzasków [23]. W wyniku tych badań ustalono postać rozkładów prawdopodobieństwa przekroczeń dla wartości quasi-szczytowej i określono zależność pomiędzy parametrami tych rozkładów a wartościami poziomów tła i krótkotrwałych impulsów zakłóceń. Wyznaczono także średni czas trwania impulsów na różnych poziomach i średnią częstotliwość ich powtarzania. Badania te pozwoliły zorientować się również w wielkości zmian poziomu zakłóceń zachodzących w ciągu doby.

Po ustaleniu postaci zakłóceń stało się możliwe przetransponowanie podstawowej grupy wyników z godzin 16—20 otrzymanych dla tła i krótkotrwałych impulsów zakłóceń na parametry rozkładu prawdopodobieństwa przekroczeń. Tą drogą uzyskano jednolity opis zakłóceń za pomocą parametrów rozkładu dla całości otrzymanych wyników.

---

<sup>2)</sup> za tło zakłóceń przyjmowano względnie stabilne wskazania miernika utrzymujące się przez okres kilkunastu sekund, natomiast za impulsy przyjmowano krótkotrwałe rzadko występujące przyrosty wskazań miernika znacznie przekraczające tło zakłóceń.

Odrębne badania przeprowadzono także we Wrocławiu i w 6 okolicznych miastach małych i średnich. Celem ich było ustalenie różnic w poziomie zakłóceń pomiędzy dachem i ulicą oraz wpływie na jej wartość rodzaju zabudowy i częstotliwości. Potrzeba tych badań wynikała z faktu, że plany sieci stacji radiofonicznych opracowuje się przy założeniu stosowania odbiorników stacjonarnych pracujących z antenami zainstalowanymi na dachach budynków. Z tego względu wymagana jest znajomość poziomu zakłóceń w miejscach ich instalowania. W wytypowanych miastach do obserwacji wybrano 172 miejsca. W miejscach tych mierzono składową pionową natężenia pola jednocześnie na dachu budynku i na ulicy przed budynkiem. Obserwacje te prowadzono na częstotliwościach 0,5, 1,6 i 2,8 MHz.

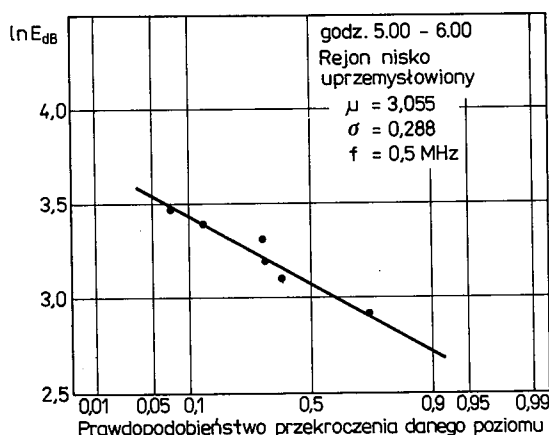
### 3.2. WYNIKI POMIARÓW

#### 3.2.1. Zależności czasowe

Na podstawie danych pomiarowych z 18 miast, wzmiankowanych powyżej, wyznaczono empiryczne rozkłady prawdopodobieństwa przekroczeń dla wartości quasi-szczytowej<sup>3)</sup>. Rozkłady te dla przykładowo wybranych godzin pokazano na rys. 6 i 7. Stwierdzono, że mogą one być aproksymowane rozkładem logarytmiczno-normalnym albo wykładniczym jeśli przyjmiemy zmienną losową  $X$  równą wartości natężenia pola w dB  $> 1 \mu\text{V/m}$ . A więc prawdopodobieństwo  $P(X > x)$  że zmienna losowa  $X$  będzie większa od  $x$  wyrazi się wzorem

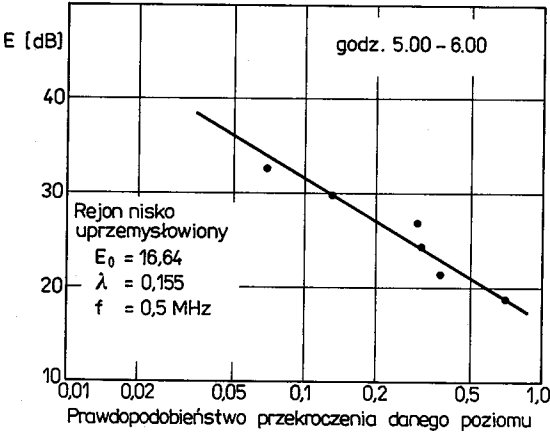
$$P(X > x) = 1 - F(x), \quad (16)$$

w którym  $F(x)$  jest dystrybuantą rozkładu logarytmiczno-normalnego o wartości oczekiwanej  $\mu$  i wariancji  $\sigma^2$ , lub jeśli zmienna  $X$  ma rozkład wykładniczy



Rys. 6. Rozkład logarytmiczno-normalny prawdopodobieństwa przekroczeń dla wartości quasi-szczytowej

<sup>3)</sup> w małych interwałach czasowych rzędu pojedynczych minut.



Rys. 7. Rozkład wykładniczy prawdopodobieństwa przekroczeń dla wartości quasi-szczytowej

$$P(X > x) = \exp[-\lambda(E_0 - x)] \tag{17}$$

gdzie  $\lambda$ ,  $E_0$  są parametrami rozkładu.

Po określeniu postaci rozkładów empirycznych dokonano oceny prawdopodobieństwa przekroczeń tła i krótkotrwałych impulsów zakłóceń, korzystając z wyników jednoczesnego pomiaru czasu przekroczeń i wartości obu tych wielkości (tablica 1). Pozwoliło to określić zależność pomiędzy parametrami rozkładów a wartościami poziomów tła  $E_t$  i krótkotrwałych impulsów  $E_i$ .

Tablica 1  
Ocena prawdopodobieństwa przekroczeń tła i krótkotrwałych impulsów zakłóceń

| Rozkład                | Tło zakłóceń | Krótkotrwałe impulsy zakłóceń |
|------------------------|--------------|-------------------------------|
| logarytmiczno-normalny | 0,877        | 0,057                         |
| wykładniczy            | 0,931        | 0,060                         |

Dla rozkładu logarytmiczno-normalnego otrzymano następujące wyrażenia

$$\mu = 0,577 \ln E_t + 0,423 \ln E_i \tag{18a}$$

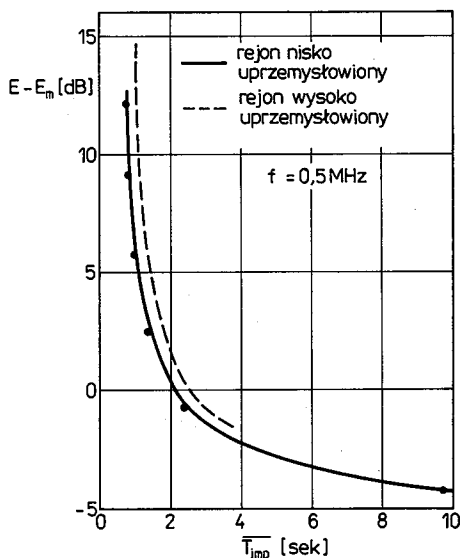
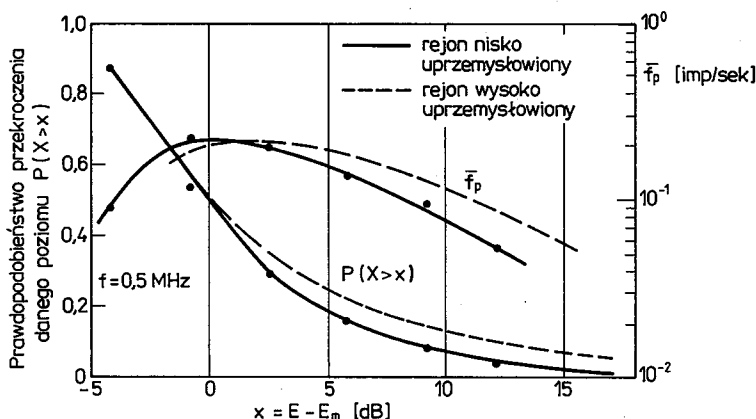
$$\sigma = 0,365 \ln(E_i/E_t), \tag{18b}$$

a dla rozkładu wykładniczego

$$E_0 = 1,0261 E_t - 0,0261 E_i \tag{19a}$$

$$1/\lambda = 0,365 (E_i - E_t), \tag{19b}$$

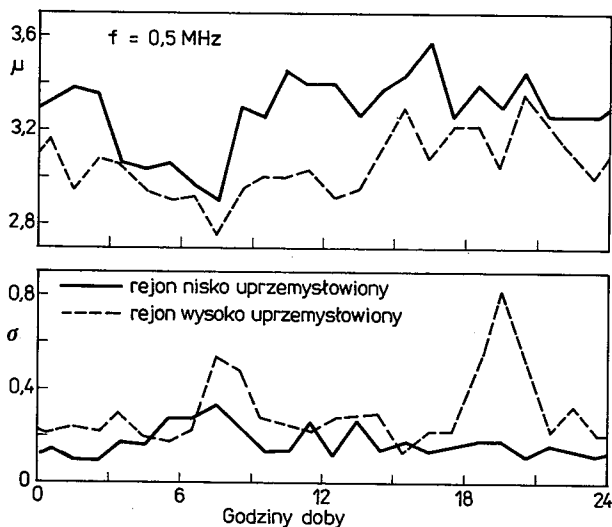
W następnej kolejności wyznaczono średnią ilość przekroczeń na jednostkę czasu  $\bar{f}_p$  i średni czas trwania impulsów  $\bar{T}_{imp}$  w zależności od poziomu zakłóceń odniesionego do

Rys. 8. Średni czas trwania impulsów na różnych poziomach.  $E_m$ -medianaRys. 9. Średnia częstotliwość powtarzania impulsów oraz rozkład prawdopodobieństwa przekroczeń.  $E_m$ -mediana

mediany ( $E_m$ ). Przebiegi te oddzielnie dla obu badanych rejonów naniesiono na rys. 8 i 9. Na rys. 9 pokazano również rozkłady prawdopodobieństwa przekroczeń.

Jak widać, różnice w przebiegu krzywych są nieznaczne. W obu wypadkach częstotliwość powtarzania nie przekracza 0,22 imp/s., przy czasie ich trwania rzędu pojedynczych sekund. Występują natomiast różnice w procencie impulsów o czasie trwania większym od 200 ms i tak dla rejonu nisko uprzemysłowionego ilość ich stanowi około 80%, a dla rejonu wysoko uprzemysłowionego około 90%.

Dobowe zmiany parametrów rozkładu prawdopodobieństwa przekroczeń ( $\mu$ ,  $\sigma$ ) dla dni powszednich pokazano na rys. 10. W obu rejonach pojawiają się wyraźne ekstrema. Najniższy poziom zakłóceń zaobserwowano około godziny 7 rano, a



Rys. 10. Przebieg dobowy zakłóceń w dni powszednie

najwyższy w godzinach popołudniowych. Pomiary przeprowadzone w niedziele obejmowały jedynie niewielki wycinek doby i były mniej liczne, w związku z tym nie pozwalają na wyciągnięcie miarodajnych wniosków. Na ich podstawie (tablica 2) można jedynie zorientować się w występujących poziomach, które jak się wydaje mogą być w pewnych wypadkach wyższe niż w dni powszednie.

W toku dalszej analizy zbudowano dystrybuanty empiryczne jednogodzinnych wartości średnich  $\mu$  i odchyłeń standardowych  $\sigma$  z okresu całej doby, łącznie dla obu badanych rejonów. Przy ich budowie wartości  $\mu$  i  $\sigma$  odniesiono do średnich z okresu czasu 16—20. Stwierdzono również, że rozkłady te są rozkładami normalnymi o parametrach podanych w tablicy 3.

W początkowym okresie badań prowadzono obserwacje poziomu zakłóceń w jednym punkcie Wrocławia przez 18 miesięcy na częstotliwości 0,5 MHz. Pomiary wykonywano codziennie, wykluczając niedziele i dni świąteczne, w godzinach między

Tablica 2

Wartość  $\mu$  i  $\sigma$  w przedziałach jednogodzinowych w niedziele

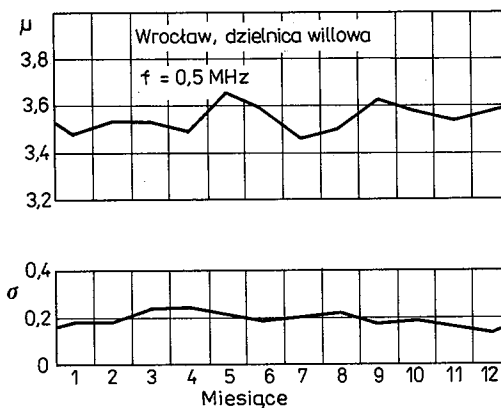
| Wyszczególnienie             | w przedziałach godzinowych |       |       |       |       |       |
|------------------------------|----------------------------|-------|-------|-------|-------|-------|
|                              | 9—10                       | 10—11 | 11—12 | 15—16 | 16—17 | 17—18 |
| rejon nisko uprzemysłowiony  |                            |       |       |       |       |       |
| $\mu$                        | 3,062                      | 3,236 | 3,245 | 3,238 | 3,271 | 3,268 |
| $\sigma$                     | 0,282                      | 0,280 | 0,273 | 0,277 | 0,264 | 0,244 |
| rejon wysoko uprzemysłowiony |                            |       |       |       |       |       |
| $\mu$                        | 3,528                      | 3,403 | 3,312 | 3,376 | 3,312 | 3,360 |
| $\sigma$                     | 0,163                      | 0,288 | 0,256 | 0,215 | 0,149 | 0,232 |

7 i 14 codziennie w innym przedziale godzinowym z zastosowaniem zmian rotacyjnych. Uzyskane z tych pomiarów wyniki zmienności rocznej pokazano na rys. 11. Jak widać, zachodzące w ciągu roku zmiany parametrów ( $\mu, \sigma$ ) są bardzo wyraźne. Największą wartość poziomu zakłóceń zaobserwowano w maju, najmniejszą w lipcu. Obraz ten w innych rejonach kraju i innych obszarach niż badany może być nieco odmienny. Na zmiany parametrów ( $\mu, \sigma$ ) w ciągu roku mogą mieć wpływ wielkość miasta, rodzaj gospodarki i zwyczaje ludności.

Tablica 3

Wartości średnie i odchylenia standardowe jednogodzinnych wartości  
 $\mu — \mu_{16-20}$  i  $\sigma — \sigma_{16-20}$  z okresu całej doby

| Wyszczególnienie          | Wartość średnia | Odchylenie standardowe |
|---------------------------|-----------------|------------------------|
| $\mu — \mu_{16-20}$       | −0,089          | 0,191                  |
| $\sigma — \sigma_{16-20}$ | −0,061          | 0,129                  |



Rys. 11. Przebieg roczny zakłóceń dla godziny 7—14

### 3.2.2. Wpływ charakteru i rodzaju zabudowy

Na wartość poziomu zakłóceń w mieście, jak wykazały rozważania teoretyczne przeprowadzone w punkcie 2, największy wpływ wywiera najbliższe otoczenie miejsca obserwacji. Z tego względu w badaniach szczególną uwagę poświęcono zabudowie miasta. We Wrocławiu, jak w każdym większym mieście, wyróżnić można strefy różniące się od innych strukturą przestrzenną, intensywnością zabudowy oraz rejonami o jednolitym rodzaju zabudowy.

Wyniki obserwacji otrzymane dla stref urbanistycznych Wrocławia zestawiono w tablicy 4. Podane wartości przedstawiają parametry rozkładu prawdopodobieństwa przekroczeń ( $\mu, \sigma$ ) dla mediany zmienności przestrzennej. W tablicy tej zestawiono

Tablica 4

Wartości  $\mu$  i  $\sigma$  przy grupowaniu wyników według stref zabudowy

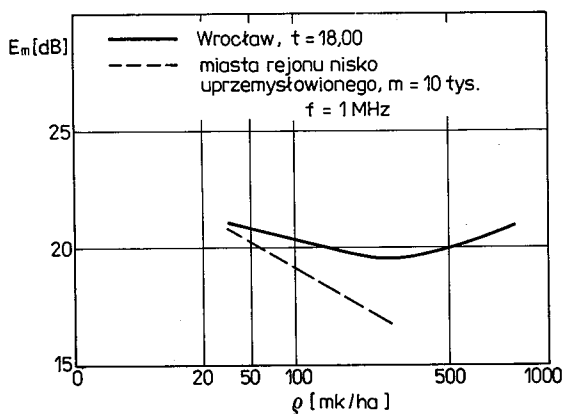
| Strefy zabudowy          | $\bar{Q}$<br>mk/ha | częstotliwość w MHz |          |       |          |       |          |       |          |       |          |
|--------------------------|--------------------|---------------------|----------|-------|----------|-------|----------|-------|----------|-------|----------|
|                          |                    | 0,5                 |          | 1,35  |          | 1,6   |          | 1,8   |          |       |          |
|                          |                    | $\mu$               | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| Wrocław                  |                    |                     |          |       |          |       |          |       |          |       |          |
| śródmiejska              | 555                | 3,258               | 0,229    | 3,143 | 0,227    | —     | —        | 2,651 | —        | 2,651 | 0,345    |
| osiedli śródmiejskich    | 122                | 3,237               | 0,249    | 3,078 | 0,215    | —     | —        | 2,616 | —        | 2,616 | 0,375    |
| osiedli podmiejskich     | 62                 | 3,283               | 0,281    | 2,970 | 0,292    | —     | —        | 2,631 | —        | 2,631 | 0,445    |
| satelitów małych         | 78                 | 3,233               | 0,190    | 2,949 | 0,228    | —     | —        | 2,612 | —        | 2,612 | 0,301    |
| o wysokiej intensywności | 555                | 3,263               | 0,230    | 3,141 | 0,227    | —     | —        | 2,651 | —        | 2,651 | 0,347    |
| o niskiej intensywności  | 97                 | 3,232               | 0,231    | 3,025 | 0,217    | —     | —        | 2,602 | —        | 2,602 | 0,368    |
| zagrodowa                | 51                 | 3,312               | 0,265    | 3,001 | 0,295    | —     | —        | 2,784 | —        | 2,784 | 0,307    |
| miasta pozostałe         |                    |                     |          |       |          |       |          |       |          |       |          |
| o wysokiej intensywności | 344                | 3,178               | 0,197    | —     | —        | 2,750 | 0,289    | —     | —        | —     | —        |
| o niskiej intensywności  | 136                | 3,181               | 0,199    | —     | —        | 2,806 | 0,263    | —     | —        | —     | —        |
| zagrodowa                | 70                 | 3,246               | 0,220    | —     | —        | 2,889 | 0,287    | —     | —        | —     | —        |

również wyniki łączne dla miast pozostałych. Przez  $\bar{q}$  oznaczono średnią dla strefy gęstość zaludnienia terenów mieszkaniowych netto w mk/ha. Śródmieście Wrocławia stanowi część miasta najgęściej zabudowaną, o dużym skupieniu usług ogólnomiejskich, z mieszaniną budynków użyteczności publicznej i mieszkaniowych. W skład jego wchodzi: osiedle historyczne, dzielnice o zabudowie z XIX w. oraz dzielnice o zabudowie nowej średniowysokiej i wysokiej.

Osiedla śródmiejskie tworzą strefę zewnętrzną. Przeważa tu zabudowa willowa i blokowa z lat trzydziestych. Całość otoczona jest strefą osiedli podmiejskich, zawierającą osiedla o zabudowie zagrodowej i osiedla mające charakter mieszany, w których przeplata się zabudowa willowa z zabudową zagrodową. Na krańcach miasta znajdują się osiedla typu tzw. satelitów małych, a które przed włączeniem ich w granice administracyjne Wrocławia miały osobny status miejski. Niektóre z tych osiedli posiadają dobrze zachowane ośrodki historyczne, na obrzeżu których występuje zarówno zabudowa blokowa jak i jednorodzinna.

Strefa o wysokiej intensywności zabudowy obejmuje budynki wielorodzinne o czterech i większej liczbie kondygnacji, zaś o niskiej intensywności budynki wielorodzinne i jednorodzinne do trzech kondygnacji [20].

Na rys. 12 pokazano zależność mediany poziomu zakłóceń od gęstości zaludnienia dla Wrocławia i miast rejonu nisko uprzemysłowionego. Wykreślono je na podstawie



Rys. 12. Zależność mediany poziomu zakłóceń od gęstości zaludnienia terenów mieszkaniowych

równań regresji otrzymanych z badań. We Wrocławiu jak można zaobserwować poziom zakłóceń ze wzrostem gęstości zaludnienia naprzód spada a następnie powyżej 300 mk/ha, gdzie występuje nadwyżka miejsc pracy w stosunku do liczby ludności, zaczyna wzrastać. Tłumaczyć można to tym, że poziom zakłóceń w mieście jest wynikiem dwu przeciwstawnych tendencji — wpływu tłumienia wnoszonego przez zabudowę i wpływu wzrostu źródeł zakłóceń. W miastach rejonu nisko uprzemysłowionego w których jak można sądzić liczba źródeł zakłóceń przypadająca na mieszkańca jest stała, poziom zakłóceń ze wzrostem gęstości zaludnienia spada.

## 3.2.3. Różnice w natężeniu pola zakłóceń między dachem a ulicą

Wyniki pomiarów, jakie zebrano na ulicy i dachach budynków, przeliczono na parametry rozkładu prawdopodobieństwa przekroczeń  $(\mu, \sigma)$ . Następnie dla par wyników z jednocześnie przeprowadzonych obserwacji, wyznaczono różnice wartości średnich  $\Delta\mu_{d-u} = \mu_d - \mu_u$  i różnice odchyłeń standardowych.  $\Delta\sigma_{d-u} = \sigma_d - \sigma_u$ . Uzyskane z tych przeliczeń wartości poddano analizie regresji wielorakiej. Badano zależność  $\Delta\mu_{d-u}$  i  $\Delta\sigma_{d-u}$  od częstotliwości i rodzaju zabudowy. W wyniku przeprowadzonych obliczeń otrzymano następujące postacie równań.

$$\Delta\mu_{d-u} = -0,210 - 0,0889 \cdot \lg f - 0,03473 \cdot \varrho^{1/3}, \quad (20a)$$

$$\Delta\sigma_{d-u} = 0,0236 + 0,0345 \cdot \lg f + 0,001989 \cdot \varrho^{1/3}, \quad (20b)$$

w których  $f$  jest częstotliwością w MHz,  $\varrho$  jest gęstością zaludnienia terenów mieszkaniowych netto w mk/ha. Poza współczynnikami regresji oszacowano także odchylenie standardowe reszt  $S_u$  i współczynnik korelacji  $R$ . Dla rów. (20a) otrzymano  $S_u = 0,174$  i  $R = 0,642$ , a dla rów. (20b)  $S_u = 0,0408$  i  $R = 0,300$ . Dane te wskazują, iż wpływ częstotliwości jak i rodzaju zabudowy na wielkość różnic  $\Delta\mu_{d-u}$  i  $\Delta\sigma_{d-u}$  jest statystycznie istotny.

## 3.2.4. Wpływ wielkości miasta i stopnia jego uprzemysłowienia

Indywidualne wyniki pomiarów z godzin 16—20, otrzymane dla tła i krótkotrwałych impulsów zakłóceń, przeliczono na parametry rozkładu przekroczeń  $(\mu, \sigma)$ . Następnie, uzyskane z tych przeliczeń wielkości poddano analizie regresji wielorakiej. Dla wartości średniej  $\mu$  i odchylenia standardowego  $\sigma$  przyjęto model matematyczny o tej samej postaci

$$y = a_0 + a_1 \varrho^{1/3} + a_2 \lg f + a_3 t + a_4 m + a_5 p, \quad (21)$$

w którym

$y$  — oczekiwana wartość średnia  $\mu$  lub odchylenie standardowe  $\sigma$ ,

$\varrho$  — gęstość zaludnienia terenów mieszkaniowych netto w mk/ha,

$f$  — częstotliwość w MHz,

$t$  — czas w godzinach,

$m$  — liczba ludności miasta w tys.,

$p$  — liczba zatrudnionych w przemyśle na 1000 mieszkańców,

$a_1, a_2, \dots, a_5$  — współczynniki równania regresji,

$a_0$  — stała równania regresji.

W wyniku przeprowadzonych obliczeń otrzymano następujące równania:

$$\mu = 2,922 - 1,437 \cdot 10^{-2} \varrho^{1/3} - 0,8215 \cdot \lg f + 2,586 \cdot 10^{-4} m + 7,134 \cdot 10^{-5} p \quad (22a)$$

$$\sigma = 0,4811 + 0,2023 \cdot \lg f - 9,325 \cdot 10^{-3} t - 1,157 \cdot 10^{-4} p \quad (22b)$$

Dla rów. (22a) otrzymano  $S_u = 0,405$  i  $R = 0,46$ , zaś dla rów. (22b)  $S_u = 0,192$  i  $R = 0,28$ . Dane te wskazują na względnie dobrą aproksymację wartości empirycznych

parametrów ( $\mu, \sigma$ ), jednak wpływ czynników nie ujętych w modelu pozostaje nadal wysoki. Nie można wykluczyć, że lepsze wyniki uzyskano by przy bardziej precyzyjnym określeniu niektórych czynników. Należy jednak nadmienić, że takie podejście prowadziłoby do skomplikowania samej postaci modelu i konieczności zebrania szeregu trudno dostępnych i nie zawsze dokładnych informacji.

Równania (20) i (22) odnoszą się do pojedynczego punktu w mieście o znanej gęstości zaludnienia. Równania 22a i b dotyczą ponadto godzin 16—20 tzn. okresu czasu w jakim przeprowadzone zostały obserwacje zakłóceń w 15% miast Polski. W praktyce interesuje nas natomiast jedna odpowiednio dobrana wielkość odnosząca się do całego miasta i całej doby. Dlatego też, tak jak już wcześniej wspomniano, za miarę poziomu zakłóceń dla miasta przyjęto wartość średnią ważoną względem liczby ludności

$$\bar{\mu} = \sum_i \eta_i \mu_i, \quad \bar{\sigma} = \sum_i \eta_i \sigma_i$$

gdzie  $\eta_i$  jest wagą wyrażającą udział ludności zamieszkującej  $i$ -ty rejon urbanistyczny w całości miasta.

Powiązanie równań (22a) i (22b) z okresem całej doby uzyskano, posługując się wynikami zmienności dobowej otrzymanymi z obserwacji poziomu zakłóceń przeprowadzonej w 18 miastach rejonu nisko i wysoko uprzemysłowionego. Na podstawie tych danych wyznaczono  $\mu - \mu_{16-20}$  i  $\sigma - \sigma_{16-20}$  (tablica 3).

Po uwzględnieniu wag i danych zawartych w tablicy 3 otrzymano następujące wyrażenia odnoszące się do poziomu zakłóceń występującego: na ulicy

$$\bar{\mu}_u = 2,771 - 0,8215 \lg f - 0,01589 \lg m + 2,586 \cdot 10^{-4} m + 7,134 \cdot 10^{-5} p \quad (23a)$$

$$\bar{\sigma}_u^2 = (0,2522 + 0,2023 \lg f - 1,157 \cdot 10^{-4} p)^2 + 0,05313 \quad (23b)$$

na dachu budynku

$$\bar{\mu}_d = 2,413 - 0,0543 \lg m - 0,9104 \lg f + 2,586 \cdot 10^{-4} m + 7,134 \cdot 10^{-5} p \quad (24a)$$

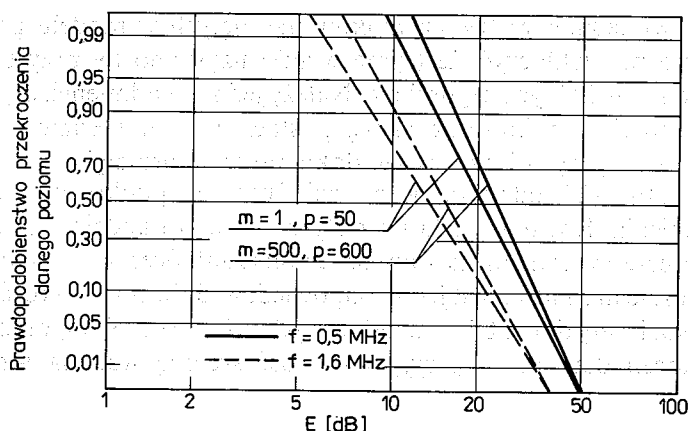
$$\bar{\sigma}_d^2 = (0,2843 + 0,2368 \lg f - 0,0022 \lg m - 1,157 \cdot 10^{-4} p)^2 + 0,05313 \quad (24b)$$

gdzie  $(\bar{\mu}_u, \bar{\sigma}_u)$  i  $(\bar{\mu}_d, \bar{\sigma}_d)$  są wartościami przekraczanymi u 50% mieszkańców i odnoszą się do całej doby.

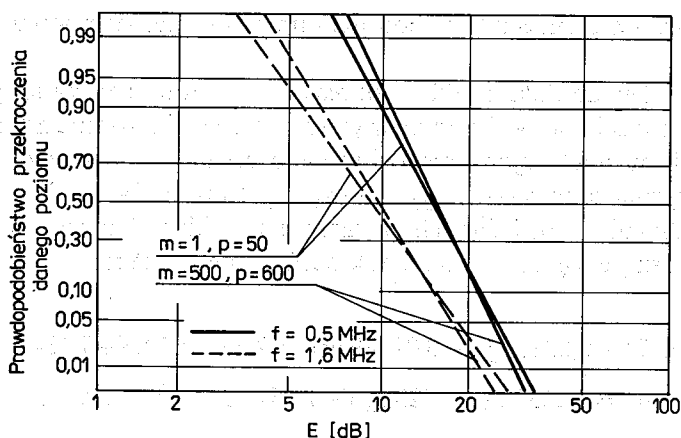
Na rys. 13 i 14 pokazano rozkłady prawdopodobieństwa przekroczeń, dla skrajnych wartości  $m$  i  $p$ , wykreślone na podstawie równań (23) i (24).

Dla zilustrowania sposobu wykorzystania otrzymanych zależności rozpatrzmy następujący przykład. Mamy miasto o liczbie ludności  $m = 35$  tys. i zatrudnieniu w przemyśle  $p = 250$  osób na 1000 mieszkańców. Należy oszacować poziom zakłóceń na dachu budynku dla częstotliwości  $f = 1$  MHz jaki jest przekraczany w 1% czasu. Z równania (24) otrzymujemy  $\bar{\mu}_d = 2,356$  i  $\bar{\sigma}_d = 0,3465$ . Te dwie wielkości określają rozkład prawdopodobieństwa przekroczeń.

Dla 1% czasu mamy  $\overline{\ln E_d} = 2,356 + 2,326 \cdot 0,3465 = 3,1619$ , skąd przewidywana wartość poziomu zakłóceń wyniesie 23,9 dB.



Rys. 13. Rozkład prawdopodobieństwa przekroczeń dla ulicy



Rys. 14. Rozkład prawdopodobieństwa przekroczeń dla dachu

## ZAKOŃCZENIE

Opracowanie zawiera materiał stanowiący podsumowanie wieloletnich badań zakłóceń przemysłowych w miastach w zakresie fal średnich radiofonicznych. Przedstawione wyniki umożliwiają ocenę poziomu zakłóceń na ulicach i dachach budynków w obrębie dzielnic mieszkaniowych. Uwzględniają one zmienność dobową i strukturę zabudowy przestrzennej miasta.

Poszczególne fragmenty badań przeprowadzone zostały w różnej skali wielkości. Przy ich uogólnieniu koniecznym było rozciągnięcie niektórych wyników na inne rejony kraju i inne częstotliwości. Takie podejście nie powinno wpłynąć na szacowaną wartość poziomu zakłóceń, zwłaszcza że rozpatruje się wąski zakres częstotliwości, w którym charakter źródeł zakłóceń nie ulega zasadniczej zmianie.

Zależności teoretyczne, jak się wydaje, mogą być pomocne przy ocenie poziomu zakłóceń dla stanu perspektywicznego w oparciu o prognozy sporządzone dla innych potrzeb jak np. prognozy rozwoju miast, wzrostu zapotrzebowania na energię elektryczną, wzrostu ilości urządzeń — potencjalnych źródeł zakłóceń itp.

## BIBLIOGRAFIA

1. CCIR Study Program 29C-1/6, *Man-made radio noise*, XVI Plenary Assembly, vol. VI, 1986
2. E.N. Skomal: *Distribution and frequency dependence of unintentionally generated man-made VHF/UHF noise in metropolitan areas*. IEEE Trans. Elektromagnetic Compatibility, cz. I, 3, 1965, cz. II 4, 1965
3. E.N. Skomal: *Analysis of the frequency dependence of man-made radio noise*, IEEE International Convention Record, 2, 1966
4. E.N. Skomal: *The range and frequency dependence of VHF-UHF man-made radio noise in and above metropolitan areas*, IEEE Trans. Vehicular Technology, 2, 1970
5. E.N. Skomal: *Distribution and frequency dependence of incidental man-made HF/VHF noise in metropolitan areas*, IEEE Trans. Elektromagnetic Compatibility, 2, 1969
6. R.T. Disney: *Estimates of man-made radio noise levels based on the office of telecommunications, its data base*. International Conference on Communications, Philadelphia, 1972
7. A.D. Spaulding: *Amplitude and time statistic of urban man-made noise*. International Conference on Communications, Montreal, 1971
8. A.U. Sheikh, J.D. Parsons: *The frequency dependence of urban man-made radio noise*, The Radio and Electric Engineer, 3, 1983
9. W.R. Lauber, J.M. Bertrand: *Man-made noise level measurement of the radio environment*. National Symp. on EMC, San Antonio, 1984
10. V.B. Jacard, R.F. Ayala: *Distance and frequency dependence of urban VHF man-made radio noise*. Inter. Conference on EMC, Surrey, 1984
11. A. Sanyal: *A survey of VHF man-made noise in the metropolitan city area of New Dehli*. Symp. on EMC, San Diego, 1986
12. Z. Rymarowicz, W. Moroń: *Obserwacje zmienności dobowej poziomu przemysłowych zakłóceń radioelektrycznych we Wrocławiu w zakresie fal średnich radiofonicznych*. Prace Instytutu Łączności, 1972 nr 1,
13. Z. Rymarowicz: *Zakłócenia radioelektryczne w zakresie fal średnich w miastach Dolnego Śląska oraz zależność od zaludnienia, wielkości i uprzemysłowienia miasta*. Praca doktorska, Instytut Łączności, 1974
14. W. Moroń, Z. Rymarowicz: *Urban man-made noise high industrialized and low industrialized regions in the medium wave range*. Internat. Symp. on EMC, Montreux 1975
15. W. Moroń, Z. Rymarowicz: *Investigation of time distribution of the composite man-made noise at a frequency 0.5 MHz in the cities of highly and sparsely industrialized regions*. Internat. Symp. on EMC, Wrocław 1976
16. W. Moroń, Z. Rymarowicz: *Zmienność dobową poziomu przemysłowych zakłóceń radioelektrycznych na częstotliwości 0.5 MHz w miastach rejonu nisko i wysoko uprzemysłowionego*. Prace Instytutu Łączności, 1977 82,
17. Z. Rymarowicz: *Zakłócenia przemysłowe w dzielnicach mieszkaniowych miasta w zakresie fal średnich radiofonicznych*. Prace Instytutu Łączności 1977 83,
18. W. Moroń, Z. Rymarowicz: *Relationship between the APD of the quasi-peak composite man-made noise and values measured with the CCIR interference meter*. Internat. Symp. on EMC, Wrocław 1978
19. R.G. Strużak, Z. Rymarowicz, W. Moroń: *Urban MF radio noise survey in Poland*. Internat. Symp. on EMC, Zurich 1983.
20. W. Czarnecki: *Planowanie miast i osiedli, tom 2 — Miejsca pracy i zamieszkania*. Warszawa: PWN, 1965

21. Z. R y m a r o w i c z: *Pomiary propagacji zakłóceń radioelektrycznych w zakresie fal średnich i krótkich na terenach mieszkaniowych*. Prace Instytutu Łączności, 1972 nr 3,
22. Z. R y m a r o w i c z: *Noise propagation in urban areas in medium and short wave range*. Internat. Symp. on EMC. Montreux 1975
23. W. R o t k i e w i c z: *Kompatybilność elektromagnetyczna w radiotechnice*. Warszawa: WKiŁ, 1978

Z. RYMAROWICZ

MAN-MADE RADIO NOISE IN THE BROADCASTING MEDIUM-FREQUENCY  
RANGE IN THE CITIES OF POLAND

S u m m a r y

The paper describes investigations of the man-made radio noise in residential quarters of towns. The research was carried out in 15% of Polish towns. Urban MF noise was experimentally investigated versus time, frequency, town size, type of and degree of industrialization. The theoretical model and experimental results are presented.

# Centrowanie parametrów technologicznych układów scalonych LSI i VLSI. Podejście eksperymentalne

MAREK J. PATYRA

*Instytut Technologii Elektronowej, Warszawa*

JAN ZABRODZKI

*Instytut Informatyki, Politechnika Warszawska*

*Otrzymano 1989.04.13*

*Autoryzowano do druku 1989.07.15*

Jednym z elementów procesu projektowania układów scalonych jest zapewnienie odpowiednio dużego uzysku produkcyjnego. W artykule rozpatrywany jest problem optymalizacji wielkości uzysku produkcyjnego poprzez właściwy dobór parametrów technologicznych z uwzględnieniem rzeczywistych warunków produkcyjnych w danej linii technologicznej. Omawiany problem określa się jako tzw. centrowanie technologiczne. W pracy zaproponowano eksperymentalno-heurystyczną metodę poszukiwania rozwiązania zadania centrowania technologicznego dla układów scalonych o dowolnej skali integracji. Metoda umożliwia praktyczną weryfikację projektu układu scalonego otrzymanego na drodze analityczno-symulacyjnej oraz optymalizację tego projektu w warunkach rzeczywistej linii technologicznej. W pracy przedstawiono wyniki zastosowania metody w praktyce produkcyjnej. Uzyskane rezultaty w pełni potwierdzają celowość stosowania metody w praktyce.

## 1. WSTĘP

Rozwojowi technologii układów scalonych towarzyszy ciągły rozwój metod projektowania i testowania tych układów. Elementem łączącym obie te dziedziny jest praktyczna realizacja opracowanego projektu. Jednym z podstawowych problemów w fazie realizacji zaprojektowanego układu scalonego jest zapewnienie odpowiednio dużego — z punktu widzenia opłacalności produkcji — uzysku produkcyjnego. Projekt układu scalonego zrealizowany metodami analityczno-symulacyjnymi powinien po wdrożeniu do produkcji w określonej linii technologicznej, gwarantować odpowiednio duży poziom uzysku technologicznego rozumianego jako stosunek liczby struktur spełniających wymagania określone dla pomiarów ostrzowych, do całkowitej liczby wytworzonych struktur danego układu.

Wielkość uzysku technologicznego jest uzależniona m.in. od różnego rodzaju defektów występujących na płytkach krzemowych. Defekty te dzieli się często na dwie klasy: defekty punktowe i defekty parametryczne [1]. Do defektów punktowych zalicza się te, które powodują skokową zmianę wartości wybranego parametru

elektrofizycznego na powierzchni zajmowanej przez co najwyżej dwie sąsiednie struktury. Defekty te są najczęściej skutkiem występowania lokalnych zanieczyszczeń materiału płytek, uszkodzeń mechanicznych płytki, lokalnych wadliwości struktury wytwarzanych warstw, defektów masek itp. Defekty parametryczne związane są z quasiciągłymi zmianami wartości parametrów elektrofizycznych elementów struktur scalonych, będących następstwem rozrzutów wartości parametrów technologicznych. Przekroczenie przez wartości tych parametrów, dopuszczalnych granic określonych w projekcie objawia się w postaci wadliwego funkcjonowania układu — w efekcie testowania mierzone parametry użytkowe układu mają wartości spoza przedziałów dopuszczalnych wartości określonych założeniami projektu.

Optymalizacja uzysku związanego z występowaniem defektów może i powinna odbywać się metodami technologicznymi (np. zaostrenie reżimu technologicznego). Ponadto w przypadku defektów punktowych istotne efekty można uzyskać już w fazie projektowania układu i określenia parametrów linii technologicznej — tego typu działania zmierzające do optymalizacji wielkości uzysku są określane w pracy jako zadanie centrowania technologicznego.

Problemy związane z centrowaniem technologicznym poruszane były dotychczas rzadko w literaturze. Spośród dostępnych publikacji można wymienić prace związane z modelowaniem i projektowaniem statystycznym układów scalonych [2, 3].

Przegląd literatury wskazuje na brak metod, które umożliwiałyby przeprowadzenie centrowania technologicznego dla współczesnych złożonych układów wielkiej i bardzo wielkiej skali integracji. Metody symulacyjne wykorzystujące koncepcję centrowania statystycznego [4, 5], stanowią w istocie próbę implementacji znanych metod optymalizacji układów elektronicznych do rozwiązywania problemów centrowania technologicznego układów scalonych. Stwarza to konieczność rozdzielenia zadania na dwie części: określenie — na podstawie rozrzutów wartości parametrów technologicznych — wartości i rozrzutów parametrów elektrycznych elementów składowych układu oraz optymalizację modelu elektrycznego układu z uwzględnieniem obliczonych rozrzutów wartości parametrów elektrycznych (np. produkcyjnie stosowany program FABRICS [6]). O zawodności stosowania w praktyce produkcyjnej metodyki rozłącznego traktowania różnych poziomów tego samego projektu przekonała się w 1987 r. firma Intel ponosząc olbrzymie straty związane z wycofaniem i dokonaniem niezbędnych poprawek w projekcie masek 32-bitowego mikroprocesora 180386 [7].

W pracy przedstawiona jest heurystyczna metoda doświadczalna centrowania technologicznego, stanowiąca niezbędne uzupełnienie metod statystycznych w procesie projektowania układów scalonych. Idea metody opisana jest w punkcie 3, zaś sama metoda w punkcie 7. W punktach 4, 5, 6 omówiono zagadnienia pomocnicze, niezbędne dla wyjaśnienia metody. Punkt 8 poświęcony jest omówieniu rezultatów stosowania proponowanej w pracy metody w praktyce projektowej.

## 2. UZYSK TECHNOLOGICZNY A PROCES PROJEKTOWANIA

W procesie produkcji układów scalonych część wytworzonych układów nie spełnia wymaganych założeń projektowych. Istnieje wiele przyczyn, które powodują, że uzysk jest zawsze mniejszy od jedności. Przyczyny te ujawniają się na różnych etapach produkcji. W niniejszym opracowaniu rozpatrujemy jedynie element składowy całkowitego uzysku produkcyjnego, związany z wytwarzaniem struktur, zwany uzyskiem technologicznym. Wartość uzysku technologicznego wyznaczana jest na podstawie wyników pomiarów ostrzowych struktur wytworzonego układu.

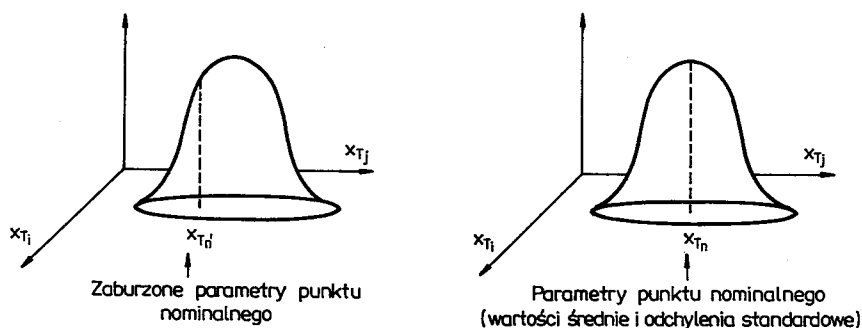
Wśród przyczyn ograniczających wielkość uzysku technologicznego wyróżnia się dwie grupy. Pierwsza z nich związana jest z występowaniem defektów punktowych, zaś druga z występowaniem defektów parametrycznych [1].

W procesie projektowania istnieje możliwość przeprowadzenia optymalizacji, która prowadzi do zmniejszenia czułości układu na defekty parametryczne. Jest to związane z takim doбором parametrów technologiczno-konstrukcyjnych (wymiały tranzystorów, grubość tlenku bramkowego tranzystora MOS, poziom domieszkowania elektrod tranzystorów, dawki implantacji domieszek, głębokości złącz, rezystancje obszarów domieszkowanych itp.), by dla danych rozrzutów statystycznych parametrów określonej linii technologicznej, zminimalizować wpływ defektów parametrycznych na wielkość uzysku technologicznego. Problem optymalizacji uzysku technologicznego jest uwzględniany w procesie projektowania układów scalonych na etapie tzw. centrowania statystycznego.

Idea metod centrowania statystycznego jest wynikiem prób adaptacji metod optymalizacji uzysku układów elektronicznych konstruowanych z elementów dyskretnych. Cechą charakterystyczną stosowania tych metod było rozdzielenie dwóch faz: fazy wytwarzania elementów i fazy wytwarzania układów z wykorzystaniem tych elementów. Proces optymalizacji związany był wyłącznie z tą drugą fazą. W przypadku układów scalonych ich struktury są wytwarzane równocześnie wraz z wytwarzaniem jej elementów składowych stąd proces projektowania powinien uwzględniać relacje między parametrami wyjściowymi układu charakteryzującymi jego cechy użytkowe a parametrami technologicznymi. W praktyce nadal jednak wykorzystuje się podejście klasyczne.

Proces optymalizacji uzysku układu scalonego prowadzi się z reguły na poziomie modelu elektrycznego, przy założonych rozrzutach parametrów technologicznych. Na sytuację taką składa się wiele przyczyn, z których zasadniczą jest brak modeli wiążących parametry użytkowe układu z parametrami technologicznymi — przynajmniej dla układów o większym stopniu scalenia niż pojedyncza bramka logiczna. Z drugiej strony dostępnych jest wiele dobrych i efektywnych programów symulacji na poziomie schematu elektrycznego (np. SPICE), które dopuszczają jednocześnie możliwość optymalizacji statystycznej. Wreszcie programy te umożliwiają w praktyce przeprowadzenie optymalizacji jedynie dla układów o niewielkiej złożoności, co oznacza, że w przypadku układów dużej skali integracji optymalizuje się ich fragmenty. Pełna symulacja współczesnych układów VLSI mogłaby trwać tysiące godzin. W praktyce po opracowaniu procesów technologicznych w danej linii uznaje się, że punkt

nominalny i dopuszczalne rozrzuty parametrów technologicznych są znane i ustalone niezależnie od rodzaju wytwarzanego układu. Pozwala to wyznaczyć rozrzuty wartości parametrów modelu elektrycznego  $X_E$  i z kolei prowadzić optymalizację statystyczną układu na poziomie modelu elektrycznego.



Rys. 1. Schematyczna ilustracja idei centrowania statystycznego

Przedstawiony wyżej stosowany w praktyce sposób projektowania układu scalonego nie pozwala na optymalizację uzysku w funkcji parametrów technologicznych (rys. 1). Sytuację można poprawić dopuszczając możliwość zmiany wartości nominalnych parametrów technologicznych już po zakończeniu procesu projektowania traktując wynik procesu projektowania jako punkt początkowy dla końcowej optymalizacji układu z uwzględnieniem rzeczywistych relacji między parametrami linii technologicznej a parametrami układu scalonego. W fazie przygotowania produkcji, można na drodze eksperymentalnej poszukiwać optymalnego punktu nominalnego z punktu widzenia uzysku. Czyni się to najczęściej metodą prób i błędów.

Teżą niniejszej pracy jest to, że etap optymalizacji eksperymentalnej prowadzonej w sposób planowy, powinien stać się integralną częścią fazy przygotowania produkcji seryjnej danego układu. Za tą tezę przemawiają następujące przesłanki. Optymalna kombinacja wartości parametrów technologicznych zależy od typu produkowanego układu, bez względu na technologię jego wytwarzania oraz od czynników związanych z konkretną linią technologiczną. Dla układów dużej skali integracji, przy braku modeli technologicznych, metod obliczania i ograniczonych możliwości symulacji, kombinacja wartości parametrów technologicznych jaką można uzyskać w procesie projektowania nie zawsze jest optymalna ze względu na uzysk. Efekty ekonomiczne związane z optymalizacją uzysku całkowicie uzasadniają poniesienie pewnych nakładów na wczesnym etapie produkcji, tym bardziej, że optymalizacja uzysku na drodze eksperymentalnej nie wiąże się z dużymi nakładami finansowymi zaś czas jej trwania jest konkurencyjny w porównaniu z pełną lub częściową symulacją komputerową. Ograniczenie się do korzystania wyłącznie z metod symulacyjnych w procesie projektowania i wdrażania układu może stać się przyczyną kosztownego w skutkach błędu, który w najlepszym przypadku ujawni się w postaci niskiej wartości uzysku, zaś w najgorszym zmusi do zmiany projektu całego układu.

### 3. CENTROWANIE TECHNOLOGICZNE

Problem centrowania technologicznego może być sformułowany następująco. Monolityczny układ scalony charakteryzuje się zbiorem parametrów wejściowych  $X$ , zbiorem parametrów wyjściowych  $Y$  oraz wiążącą te zbiory relacją określoną przez funkcję  $F: Y = F(X)$ . Do zbioru  $n$  parametrów wejściowych zalicza się napięcia zasilające, wielkości fizyczne oddziałujące na wejścia układu, parametry technologiczne — określone zbiorem  $X_T$  (np. dawki implantacji domieszek, głębokości złącz, grubości warstw izolacyjnych itp.), parametry topograficzne (np. wymiary poszczególnych obszarów) oraz parametry elektrofizyczne (np. napięcia progowe tranzystorów MOS, współczynniki wzmocnienia tranzystorów, pojemności pasożytnicze itp.).

Zbiór  $m$  parametrów reprezentujących cechy użytkowe układu w postaci wartości napięć i prądów na wyjściach układu oraz zależności czasowe między sygnałami wejściowymi bądź sygnałami wyjściowymi a sygnałami wejściowymi tworzy zbiór parametrów wyjściowych  $Y$ .

W  $(n+m)$  wymiarowej przestrzeni parametrów  $X$  i  $Y$  można wyróżnić zbiór punktów, dla których układ scalony jest sprawny, tzn., że dla danej kombinacji wartości parametrów wejściowych dany układ spełnia swoje funkcje a wartości parametrów wyjściowych należą do dopuszczalnych przedziałów ich zmian. Zbiór ten tworzy tzw. bryłę charakteryzacyjną. Ograniczając tę bryłę do zbioru kombinacji parametrów wejściowych, dla których układ jest sprawny otrzymuje się tzw. bryłę sprawności. Jeżeli ponadto założyć się że są ustalone wartości wszystkich parametrów wejściowych poza wartościami parametrów technologicznych, to bryła sprawności ogranicza się do technologicznej bryły sprawności.

Wiążąc z wybranym punktem technologicznym rozrzuty wartości parametrów (tolerancje bądź rozkłady gęstości prawdopodobieństw), można badać wielkość uzysku wynikającego z faktu istnienia rozrzutów wartości tych parametrów. W szczególności można poszukiwać takiego położenia nominalnego punktu technologicznego, dla którego uzysk jest maksymalny. Rozwiązanie takiego zadania na drodze analitycznej lub symulacyjnej dla współczesnych układów scalonych wielkiej skali integracji nie jest obecnie całkowicie możliwe. Możliwe jest natomiast wykorzystanie metod eksperymentalnych. Wtedy zadanie centrowania sprowadza się do znalezienia optymalnego punktu nominalnego na drodze planowo ustalonej serii eksperymentów realizowanych w warunkach rzeczywistej linii technologicznej. Przez eksperyment rozumie się wytworzenie partii układu scalonego składającej się z kilku grup płytek, przy czym każda grupa jest wykonana dla innego punktu technologicznego, oraz dokonanie pomiarów ostrzowych tych płytek. Plan wykonywania eksperymentu musi być ustalony w warunkach nieznajomości zarówno bryły sprawności jak i funkcji wiążącej parametry wyjściowe z parametrami technologicznymi. Możliwe są różne strategie postępowania przy ustalaniu planów kolejnych eksperymentów. Często o wyborze danej strategii decyduje doświadczenie lub intuicja konstruktora [8].

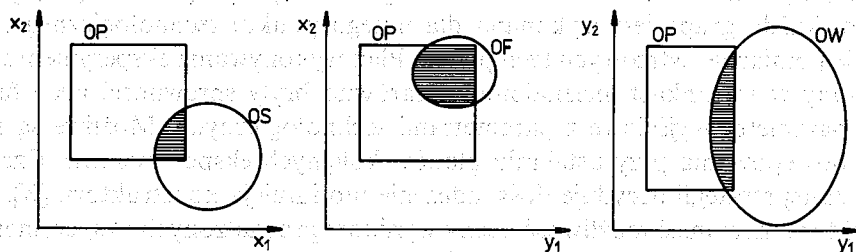
Projektant musi mieć możliwość oceny wyników prowadzonych eksperymentów. Narzucającym się kryterium oceny jest wielkość uzysku technologicznego w danym punkcie technologicznym. Jednakże stosowana w praktyce metoda wyznaczania

uzysku nie może być przyjęta jako odpowiedni wskaźnik oceny. Przy określaniu wielkości uzysku na podstawie pomiarów ostrzowych przeprowadzonych jedynie dla wartości nominalnych parametrów wejściowych nie ma możliwości oceny wpływu defektów punktowych i parametrycznych na wielkość uzysku a jak stwierdzono wyżej tylko te ostatnie są istotne z punktu widzenia procesu centrowania technologicznego. Konieczne jest zatem znalezienie innego kryterium oceny, które eliminowałoby wpływ defektów punktowych.

#### 4. WEKTOR KIERUNKOWY

W praktyce dokonuje się redukcji bryły charakterystycznej do dwuwymiarowych (trójwymiarowych) przekrojów, poprzez ustalenie wartości wszystkich parametrów poza wybranymi dwoma (trzema). Przekroje takie mogą być wyznaczone dla dwóch parametrów wejściowych, dwóch parametrów wyjściowych lub parametru wejściowego i wyjściowego. Przekroje takie nazywane są odpowiednio obszarami sprawności, obszarami pracy i obszarami funkcjonalnymi [12]. Obszary te są zbiorem punktów w odpowiednich przestrzeniach, dla których układ spełnia wymagania projektowe.

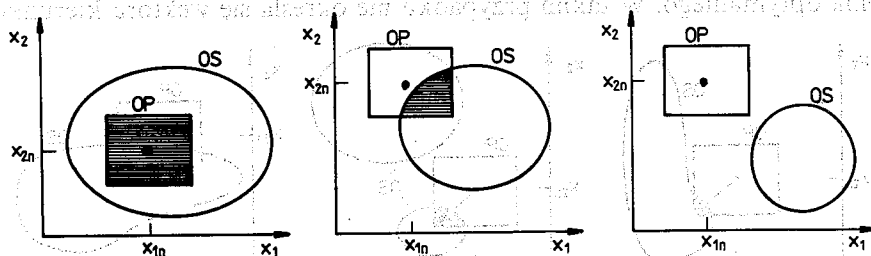
W trakcie pomiarów ostrzowych możliwe jest sprawdzenie działania układu dla zbioru punktów należących do odpowiednio dobranego okna pomiarowego określonego przez minimalne i maksymalne wartości parametrów w czasie pomiarów. Weźmy pod uwagę płaszczyznę obszaru sprawności. Przeszukując systematycznie okno pomiarowe (z przyjętym dla każdego parametru krokiem pomiarowym) można wyznaczyć zbiór punktów, dla których dany układ jest sprawny. Zbiór ten stanowi iloczyn obszaru sprawności i obszaru okna pomiarowego. Zauważmy, że jeśli wymiary okna pomiarowego są zgodne z dopuszczalnymi przedziałami zmian wartości parametrów wejściowych, to część wspólna tych obszarów określa wielkość uzysku związanego tylko z występowaniem defektów parametrycznych — część wspólna może być niepusta jedynie wtedy gdy nie występuje defekt punktowy. Wniosku takiego nie można wyciągnąć na podstawie tradycyjnie prowadzonych pomiarów ostrzowych (tylko w punkcie nominalnym). Oczywiście wniosku takiego nie można wyciągnąć również w przypadku zerowej części wspólnej okna pomiarowego i obszaru sprawności; możliwym rozwiązaniem jest wówczas zwiększenie wymiarów okna pomiarowego.



Rys. 2. Badanie części wspólnej okna pomiarowego (OP) i obszaru sprawności (OS), obszaru funkcjonalnego (OF) oraz obszaru pracy (OW)

Podobnie jak dla obszaru sprawności (OS), można postępować również w odniesieniu do obszarów pracy (OW) i obszarów funkcjonalnych (OF) (rys. 2).

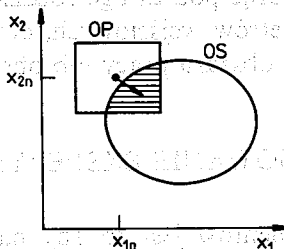
Załóżmy, że okno pomiarowe ma wymiary zgodne z dopuszczalnymi przedziałami zmian wartości parametrów. Okno pomiarowe pokrywa się w tym przypadku z prostokątem tolerancji. Możliwe relacje między prostokątem tolerancji a odpowiednim obszarem są zilustrowane schematycznie na rys. 3.



Rys. 3. Schematyczna ilustracja wzajemnych relacji obszaru sprawności i prostokąta tolerancji

Przypadek pokazany na rys. 3(a) odpowiada sytuacji optymalnej, gdy cały prostokąt tolerancji zawiera się w obszarze sprawności i z punktu widzenia uzysku (ze względu na defekty parametryczne), każda wyprodukowana struktura jest sprawna.

W przypadku zilustrowanym na rys. 3(b) wielkość uzysku jest mniejsza od jedności co oznacza, że punkt technologiczny nie jest wybrany optymalnie. Stosunek powierzchni okna pomiarowego do części wspólnej okna i obszaru sprawności wyraża wielkość odstępstwa od optimum. Stosunek ten można reprezentować za pomocą wektora zaczepionego w środku geometrycznym okna pomiarowego i skierowanego do środka ciężkości obszaru sprawności leżącego w tym obszarze. Tak zdefiniowany wektor nazywany jest wektorem kierunkowym (rys. 4).

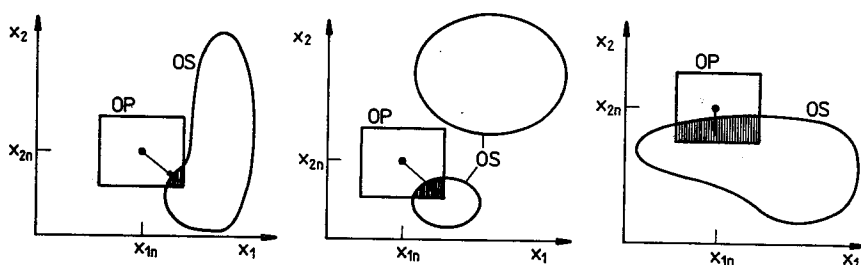


Rys. 4. Ilustracja definicji wektora kierunkowego

Znajomość wektora kierunkowego dla danego punktu technologicznego nie niesie konkretnej informacji o tym, w którym kierunku należy przesunąć punkt technologiczny w przestrzeni parametrów technologicznych, aby uzyskać pożądane zmniejszenie długości wektora kierunkowego. Wektor kierunkowy wyznaczony dla pojedynczego punktu technologicznego nie zawsze wskazuje na rzeczywisty kierunek, w którym należy przesunąć prostokąt tolerancji by zwiększyć jego część wspólną z

obszarem sprawności. Dla ilustracji na rys. 5 pokazano trzy przykłady dla niewypukłych lub niespójnych obszarów sprawności. Pożądaną informację o kierunku możliwych zmian można uzyskać analizując wektory kierunkowe dla kilku punktów technologicznych a to z kolei pozwala na planowe prowadzenie eksperymentów technologicznych.

Sytuacja pokazana na rys. 3(c) może wystąpić gdy układ jest niesprawny na skutek obecności defektu punktowego lub gdy punkt technologiczny leży bardzo daleko od położenia optymalnego. W takim przypadku nie określa się wektora kierunkowego.



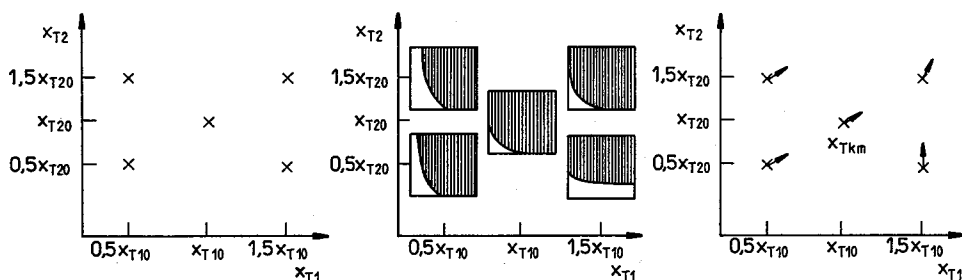
Rys. 5. Wyznaczenie wektorów kierunkowych dla przypadków obszarów wklęsłych i niespójnych

W praktyce wielkość okna pomiarowego jest większa od wielkości prostokąta tolerancji. Ułatwia to rozstrzygnięcie wątpliwości charakterystycznych dla przypadku zilustrowanego na rys. 3(c), a ponadto zmniejsza prawdopodobieństwo wystąpienia wektorów kierunkowych znacznie odbiegających od wektorów określających właściwy kierunek zmian. Stosując większe rozmiary okna pomiarowego zbiera się większą ilość informacji o badanym układzie. Wymiary okna nie mogą być jednak dowolnie zwiększane co najmniej z dwóch powodów: nie można dopuścić do zniszczenia układu w trakcie badań ani też do wystąpienia sytuacji, w której cały obszar sprawności znalazłby się w oknie pomiarowym. Wówczas, jak w przypadku pokazanym na rys. 3(c) definiowanie wektora kierunkowego traci sens. Wielkość okna pomiarowego ustala zwykle konstruktor układu biorąc pod uwagę rozdzielczość urządzenia testującego, ograniczenia wielkości parametrów wejściowych, ewentualną znajomość kształtu i wielkości obszaru sprawności, obszaru pracy lub obszaru funkcjonalnego.

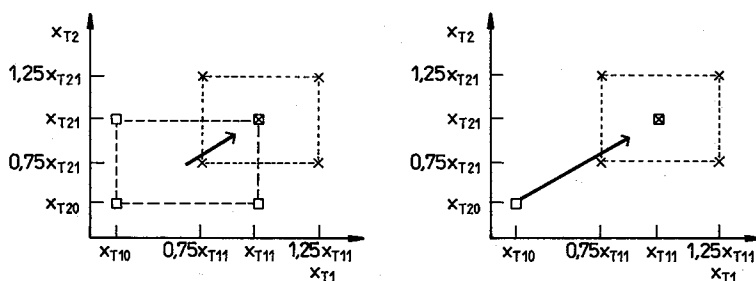
## 5. PLANOWANIE EKSPERYMENTU

Ideą planowania eksperymentów jest to aby na podstawie niewielkiej liczby wykonanych partii układów, uzyskać jak najszybciej informacje o położeniu optymalnego punktu technologicznego. Przy braku znajomości funkcji opisującej zależność parametrów wyjściowych od parametrów wejściowych można korzystać jedynie ze strategii o charakterze heurystycznym; wybór określonej strategii zależy często od doświadczenia i intuicji projektanta jak też od ograniczeń technologicznych (maksymalna liczba płytek w partii, liczba procesów technologicznych itp.) oraz ekonomicznych. Poniżej opisano dwie strategie postępowania wykorzystujące planowanie eksperymentu według planu statycznego zdeterminowanego kompletnego [8].

Założmy, że dopuszczalne są zmiany parametrów technologicznych  $X_{T1}$  i  $X_{T2}$  oraz, że znany jest początkowy punkt technologiczny o współrzędnych  $(X_{T10}, X_{T20})$ . Traktując punkt początkowy jako centralny punkt planu wykonuje się partię płytek zawierającą kilka grup płytek dla różnych kombinacji parametrów technologicznych przedstawionych na rys. 6(a). Dla każdego punktu technologicznego wyznacza się odpowiedni obszar (rys. 6(b)) oraz wektory kierunkowe, uzyskując mapę wektorów zilustrowaną na rys. 6(c).



Rys. 6. Określenie mapy wektorów kierunkowych dla pięciu grup płytek wytworzonych według planu eksperymentu technologicznego



Rys. 7. Strategie planowania eksperymentów technologicznych

W strategii pierwszej jako punkt centralny dla każdego następnego eksperymentu, przyjmuje się punkt  $X_{Tkm}$ , dla którego wektor kierunkowy ma najmniejszą wartość. Wokół tego punktu ustala się siatkę punktów, dla których wykonywane są następne grupy płytek (patrz rys. 7(a)).

W strategii drugiej dla wyznaczenia punktu centralnego planu określa się wypadkowy wektor kierunkowy, zaczepiony w punkcie początkowym (wektor zaznaczony grubą linią na rys. 7(a)). Punkt centralny planu kolejnego eksperymentu wybiera się na kierunku wskazanym przez wypadkowy wektor kierunkowy, w odległości uzależnionej od jego długości. Pozostałe punkty planów ustala się podobnie jak w strategii pierwszej (rys. 7(b)).

Jeśli kolejna faza eksperymentu nie daje pożądanego rezultatu (w postaci znalezienia punktu technologicznego dla którego okno pomiarowe leży całkowicie wewnątrz obszaru sprawności), procedurę powtarza się wyznaczając plan kolejnego eksperymentu.

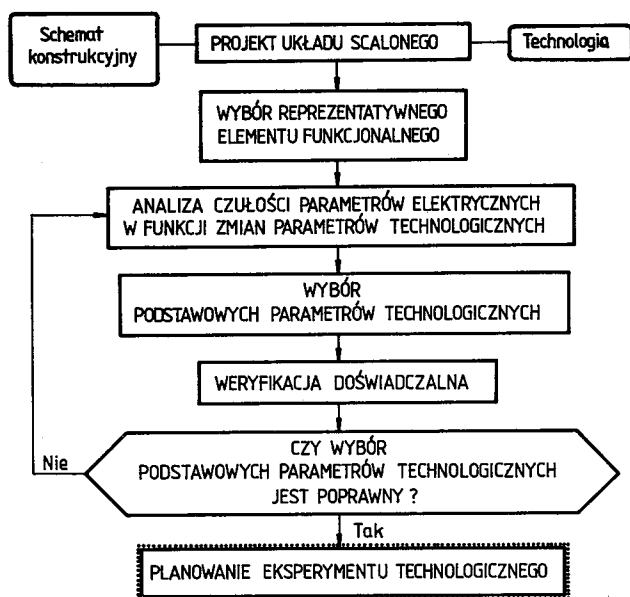
Różnica między obu przedstawionymi strategiami sprowadza się do tego, że w strategii pierwszej punkt centralny kolejnego planu jest wyznaczany niezależnie od kierunków wektorów kierunkowych, podczas gdy w strategii drugiej informacja taka jest brana pod uwagę. W praktyce może się okazać celowe wprowadzenie dodatkowych modyfikacji zmieniających odległość między punktami planu bądź siatkę punktów planu w zależności od wyników analizy poprzednich eksperymentów. Korzystna może się okazać również zmiana samego programu badań np. poprzez zastosowanie planów selekcyjnych [8]. Zmiana strategii dokonywana przez konstruktora układu opiera się również na przesłankach heurystycznych.

## 6. WYBÓR PODSTAWOWYCH PARAMETRÓW TECHNOLOGICZNYCH I PARAMETRÓW KRYTYCZNYCH

Pośród parametrów technologicznych zawsze można wyróżnić takie, których zmiany mają mniej lub bardziej istotny wpływ na własności funkcjonalne a tym samym na wartości parametrów wyjściowych układu scalonego. Ze względu na dążenie do zachowania rozsądnego kompromisu między dokładnością eksperymentu a jego kosztami, z reguły istnieje konieczność wybrania dwóch lub trzech parametrów technologicznych o największym oddziaływaniu na funkcjonowanie układu. Parametry te określa się jako podstawowe parametry technologiczne  $X_{TP}$ .

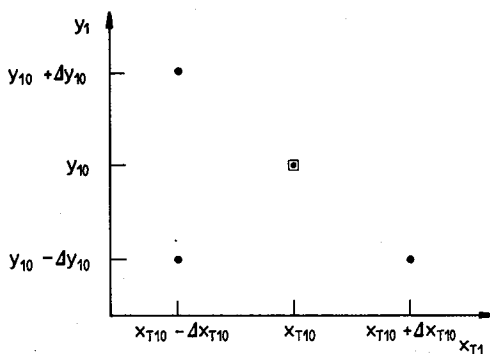
W ogólności, wyboru podstawowych parametrów technologicznych można dokonać wykorzystując znane metody badania wrażliwości układów [10]. Wymaga to jednak znajomości pełnego modelu układu scalonego uwzględniającego zależności między parametrami wyjściowymi a parametrami technologicznymi. Wobec braku podobnych modeli konieczne jest korzystanie z metod uproszczonych: symulacyjnych, doświadczalnych lub ich kombinacji. Korzystając z dostępnych metod symulacji można ograniczyć się do badania wpływu zmian parametrów technologicznych na parametry elektryczne elementów lub bloków funkcjonalnych układu [11] wykorzystując np. metody analizy czynnikowej [10]. Proponowany schemat postępowania przy określaniu podstawowych parametrów technologicznych przedstawiono na rys. 8 [13].

Do celów oceny pracy układu scalonego w ustalonym punkcie technologicznym proponowano wyżej wykorzystanie koncepcji bryły charakterystycznej. Ze względów praktycznych spośród znacznej liczby parametrów wyjściowych korzystne jest wybranie kilku (np. trzech) w stosunku do których wyznacza się odpowiednie obszary sprawności, funkcjonalne lub obszary pracy przydatne do analizy i oceny danego układu. Wybrane parametry, określane dalej jako krytyczne, powinny charakteryzować się dużą czułością na zmiany parametrów technologicznych. Dodatkowo należy zapewnić łatwość pomiarów badanych obszarów względem wybranych parametrów. Ponieważ w praktyce najwygodniej jest posługiwać się obszarami funkcjonalnymi ze względu na zestaw parametrów wejściowo-wyjściowych, należy po ustaleniu podstawowych parametrów technologicznych dokonać wyboru par para-



Rys. 8. Schemat postępowania przy określaniu podstawowych parametrów technologicznych

metrów krytycznych. Można tego dokonać doświadczalnie, mierząc struktury układów wytworzone w trakcie realizacji pierwszego eksperymentu, według przykładowego planu pokazanego na rys. 9.

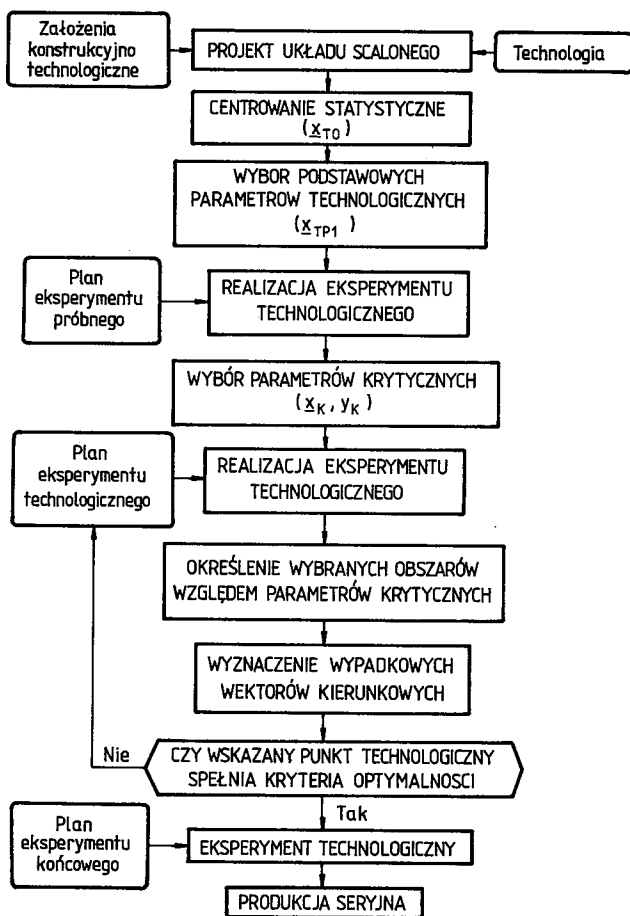


Rys. 9. Plan pierwszego eksperymentu technologicznego

## 7. METODA CENTROWANIA TECHNOLOGICZNEGO

Na rys. 10 pokazano schemat blokowy ilustrujący podstawowe kroki proponowanej metody centrowania technologicznego.

Po realizacji projektu układu scalonego następuje etap centrowania statystycznego metodami symulacyjnymi. W wyniku symulacji uzyskuje się współrzędne (parametry) punktu  $X_{T0}$  dla którego zgodnie z analizą teoretyczną możemy spodziewać się



Rys. 10. Schemat metody centrowania technologicznego

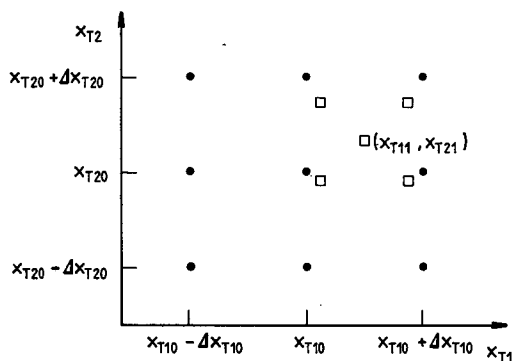
maksymalnego uzysku. Dalej następuje etap wyboru podstawowych parametrów technologicznych, a po ich wybraniu wykonuje się próbną partię układów dla kilku punktów technologicznych. Dla każdego z tych punktów wybiera się odpowiednią kombinację wartości parametrów podstawowych zgodnie z przyjętą strategią planowania eksperymentu (na przykład są to kombinacje wartości  $X_{T1}$  i  $Y_1$  — patrz rys. 10). Wartości pozostałych parametrów technologicznych są ustalone. Na podstawie analizy wybranych obszarów (sprawności, pracy lub funkcjonalnych), określonych dla wytworzonych struktur, wyznacza się pary parametrów krytycznych [10]. Następnie sprawdza się czy w którymś z punktów technologicznych zachodzi warunek optymalności (okno pomiarowe całkowicie pokryte przez badany obszar). Jeśli tak to przechodzi się do realizacji eksperymentu końcowego, w przeciwnym przypadku wykonuje się serię eksperymentów technologicznych zgodnie z przyjętą strategią.

W wyniku zastosowania dotychczasowej procedury zostaje znaleziony punkt technologiczny, dla którego okno pomiarowe zawarte jest w badanym obszarze. Jednakże z punktu widzenia rozwiązania zadania centrowania technologicznego, za

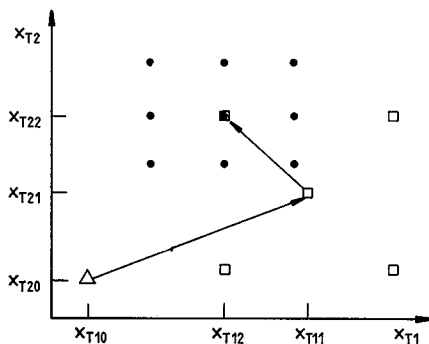
optymalny można uznać taki punkt nominalny, dla którego w każdym punkcie należącym do związanego z nim prostokąta tolerancji spełniony jest warunek zawierania okna w badanym obszarze. Dla wyznaczenia takiego punktu przeprowadza się eksperyment końcowy, przyjmując plan, którego schemat zilustrowano na rys. 11 (punkty planu oznaczone  $\square$ ). Ostatecznie za poszukiwany punkt optymalny należy przyjąć ten spośród punktów planu końcowego, dla którego spełniony jest warunek optymalności rozszerzony o wymaganie aby dla sąsiednich punktów z planu okna pomiarowe były również pokryte przez badane obszary. W przypadku niewystąpienia podobnej sytuacji należy przeprowadzić dalsze eksperymenty pomocnicze.

Prezentowana metoda centrowania technologicznego jest metodą heurystyczną i jej skuteczność zależy od aktywnego udziału konstruktora układu scalonego, który w poszczególnych etapach musi podejmować decyzje na podstawie wcześniej uzyskanych rezultatów, zwłaszcza przy ustalaniu planów kolejnych eksperymentów.

Przykładowy przebieg poszukiwania rozwiązania zadania optymalizacji z zastosowaniem metody centrowania technologicznego pokazano na rys. 12. Po ustaleniu podstawowych parametrów technologicznych oraz parametrów krytycznych, stwier-



Rys. 11. Plan końcowego eksperymentu technologicznego



Rys. 12. Przykładowy przebieg poszukiwania punktu optymalnego z zastosowaniem metody centrowania technologicznego

dzono, że wektor kierunkowy w punkcie startowym ( $X_{T10}, X_{T20}$ ) posiada bardzo dużą długość i ustalono plan pierwszego pięciopunktowego eksperymentu. Centralny punkt planu tego eksperymentu stanowi punkt ( $X_{T11}, X_{T21}$ ) zaznaczony kwadratem na rys. 12. Warunek optymalności spełniony jest tylko w jednym punkcie planu (oznaczonym krzyżykiem), który przyjęto za centralny punkt dziewięciopunktowego planu końcowego eksperymentu. Wyniki tego eksperymentu potwierdziły przypuszczenie iż punkt ( $X_{T12}, X_{T22}$ ) stanowi poszukiwany punkt optymalny.

Łącznie w omawianym przykładzie zrealizowano dwie partie układów, liczące odpowiednio: pięć i dziewięć płytek (nie licząc grupy płytek zrealizowanych dla punktu startowego). Jeśli przyjąć, że każda grupa zawierała średnio trzy płytki, to do realizacji opisanej serii eksperymentów zużyto czterdzieści pięć płytek krzemowych.

## 8. WERYFIKACJA DOŚWIADCZALNA METODY

Dla zweryfikowania przedstawionej metody i zbadania możliwości eksperymentalnego zwiększenia uzysku na drodze centrowania technologicznego, przeprowadzono optymalizację parametrów technologicznych dwóch wybranych układów MOS LSI, wytwarzanych w technologii NMOS. Pierwszy układ spełnia funkcje programowego, równoległego interfejsu danych, zaś drugi programowego szeregowego interfejsu danych w systemach mikroprocesorowych.

Wyboru podstawowych parametrów technologicznych dokonano zgodnie z procedurą, której schemat pokazano na Rys. 8. Podstawowym elementem konstrukcji elektrycznej projektowanych układów jest inwerter MOS typu ED. Korzystając z wyników opublikowanych w pracy [11] oraz z symulacji charakterystyk statycznych i dynamicznych fragmentów konstrukcji elektrycznej układów wybrano dwa podstawowe parametry technologiczne: dawki implantacji fosforu i dawki implantacji boru. Metodą analizy czynnikowej określono czułość parametrów dynamicznych inwertorów na zmiany parametrów technologicznych.

Potwierdzenie słuszności dokonanego wyboru uzyskano już podczas realizacji pierwszego eksperymentu technologicznego. Niżej omówione są implementacje metody centrowania technologicznego dla wybranych układów.

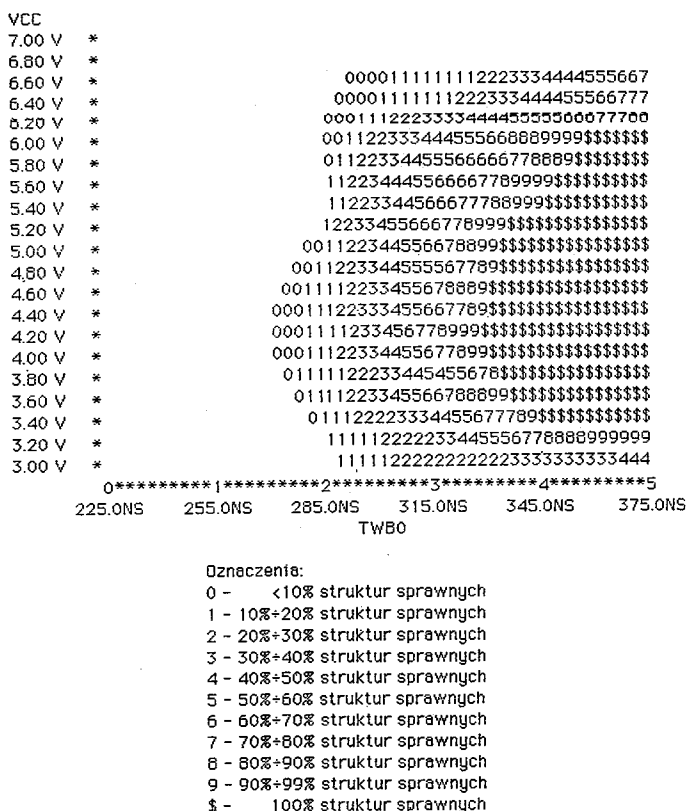
### 8.1. UKŁAD PIERWSZY

Traktując punkt technologiczny uzyskany w projekcie teoretycznym jako punkt startowy zaplanowano eksperyment dwuczynnikowy przyjmując pięćdziesięć procentowe odchylenia od wartości parametrów punktu startowego jako parametry punktów planu.

Jednocześnie dokonano wstępnej selekcji parametrów  $X$  i  $Y$  pod względem zastosowań pomiarowych. Znając konstrukcję układu zdecydowano, że parametrów krytycznych należy poszukiwać w zbiorze zawierającym napięcie zasilania  $U_{DD}$  oraz parametry dynamiczne układu (zależności czasowe między sygnałami wejściowymi i wyjściowymi). Ostatecznie wybrano około dziesięciu par parametrów; jednym z tych parametrów było zawsze napięcie zasilające. Dla wytworzonych czterech grup płytek wyznaczono obszary funkcjonalne względem par parametrów wejściowo-wyjściowych. Po określeniu wielkości wektorów kierunkowych okazało się, że dla siedmiu par parametrów zmiany długości wektorów były nieznaczne (do kilkunastu procent dla czterech badanych punktów technologicznych), natomiast dla trzech pozostałych zmiany te sięgały 75%. Do dodatkowych badań przyjęto więc te trzy pary parametrów; ostatecznie jako drugi element pary wybrano czas propagacji danej z wejścia na wyjście w jednym z trybów pracy, oznaczany dalej jako  $t_{WBO}$ . W trakcie pomiarów stosowano wstępną obróbkę danych dającą w rezultacie zbiorcze obszary funkcjonalne dla struktur z całej płytki. Przykład zbiorczego obszaru funkcjonalnego dla badanego układu pokazano na rys. 13.

W trakcie dalszego postępowania, w celu eliminacji możliwych różnic w wytwarzaniu struktur badanego układu (również układu drugiego), licznosc płytek krzemowych

w poszczególnych grupach, objętych planami eksperymentów była tak dobrana aby płytki stanowiły jedną partię technologiczną, a realizacja procesów technologicznych odbywała się wspólnie z wyjątkiem procesu, którego parametr został wybrany do optymalizacji.



Rys. 13. Przykład zbiorczego obszaru funkcjonalnego struktur badanego układu z jednej płytki

Projekt układu nie był poddawany centrowaniu statystycznemu a wartości parametrów w początkowym punkcie technologicznym stanowiły średnie arytmetyczne wartości maksymalnych i minimalnych uzyskanych z analizy układu na najgorszy przypadek.

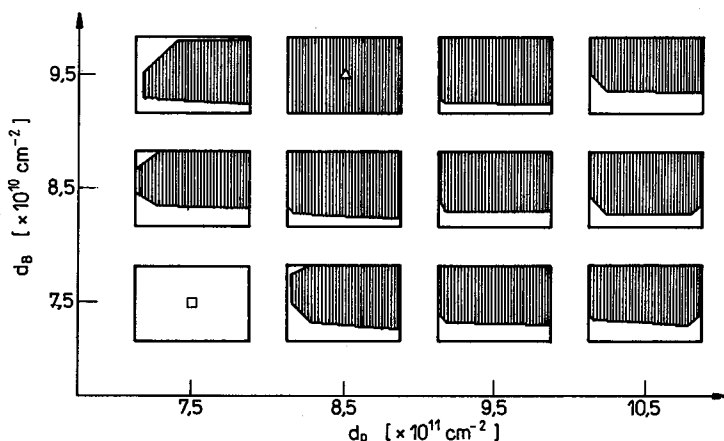
Na rysunku 14 jest pokazany obszar funkcjonalny struktury wytworzonej zgodnie z parametrami początkowego punktu technologicznego. Wektor kierunkowy jest długi co wskazuje na dużą odległość od poszukiwanego punktu optymalnego. Dla omawianego punktu uzysk miał niewielką wartość co świadczy o tym, że wyznaczenie punktu technologicznego metodą „na najgorszy przypadek”, może być obciążone bardzo dużym błędem dla układów LSI i VLSI. Zaistniała więc konieczność przeprowadzenia eksperymentu technologicznego pozwalającego w przybliżeniu zlokalizować położenie punktu nominalnego. Wyniki analizy przeprowadzonej na etapie projektu teoretycznego („na najgorszy przypadek”) pozwoliły na oszacowanie minimalnych i maksymalnych dopuszczalnych zmian wartości optymalizowanych





$x_{T2} = 8.5 \times 10^{10}$  [1/cm<sup>2</sup>]. Tym razem plan eksperymentu obejmuje dwanaście punktów technologicznych (parametr  $x_{T1}$  cztery wartości i parametr  $x_{T2}$  trzy wartości) z tą różnicą, że zmniejszono odległości między punktami planu. Na rys. 16 przedstawiono również rozmieszczenie punktów drugiego planu. Sieć punktów planu wybrano w taki sposób aby punkty planów obu eksperymentów, odpowiadających partiom 1 (■) i 2 (□) częściowo się pokrywały.

Po wytworzeniu partii struktur badanego układu i wykonaniu pomiarów stwierdzono, że grupa płytek odpowiadająca punktowi oznaczonemu □ na rys. 18 jest zrealizowana wadliwie. Dla pozostałych grup płytek uzyskano rozkład wektorów kierunkowych pokazany na rys. 16. W punkcie technologicznym oznaczonym przez Δ



Rys. 18. Mapa obszarów funkcjonalnych badanych układów wytworzonych w punktach odpowiadających planowi drugiego eksperymentu

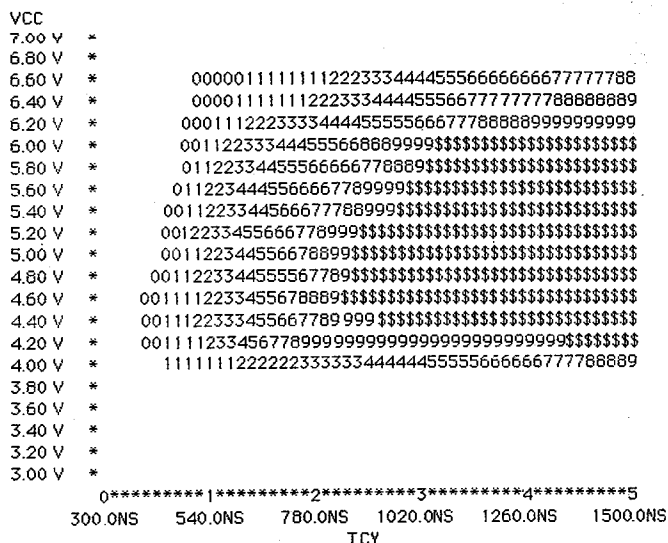
uzyskano pełne pokrycie okna pomiarowego przez badany obszar. Ponieważ wskazany punkt jest punktem brzegowym eksperymentu, dokonano uzupełnienia eksperymentu o dodatkowe cztery punkty oznaczone ● na rys. 15. Stwierdzono wzrost długości wektorów kierunkowych wraz z oddalaniem się od wskazanego punktu. W efekcie poszukiwany, w procesie centrowania technologicznego, punkt optymalny ma współrzędne:  $d_B = 9.5 \times 10^{10}$  [1/cm<sup>2</sup>] i  $d_p = 8.5 \times 10^{11}$  [1/cm<sup>2</sup>]. Słuszność dokonanego wyboru potwierdziły badania serii partii prototypowych, dla których uzyskano dużą powtarzalność mierzoną stopniem pokrycia okien pomiarowych przez badane obszary dla trzech wybranych par parametrów krytycznych. Dodatkowym faktem świadczącym o poprawności zastosowanej metody są średnie wartości uzysku dla wytwarzanych partii kilkakrotnie większe od poziomu opłacalnego uzysku.

## 8.2. UKŁAD DRUGI

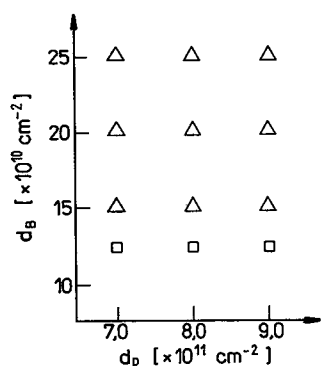
Podczas opracowania drugiego z wymienionych układów scalonych przystąpiono bezpośrednio do etapu wyboru parametrów krytycznych. Ze względu na duże podobieństwo technologiczne obu układów uznano, że podstawowe parametry

technologiczne określone dla pierwszego układu mogą być bezpośrednio wykorzystane podczas optymalizacji parametrów układu drugiego. Spodziewać się należy jedynie różnic ilościowych wynikających z odmiennych konstrukcji.

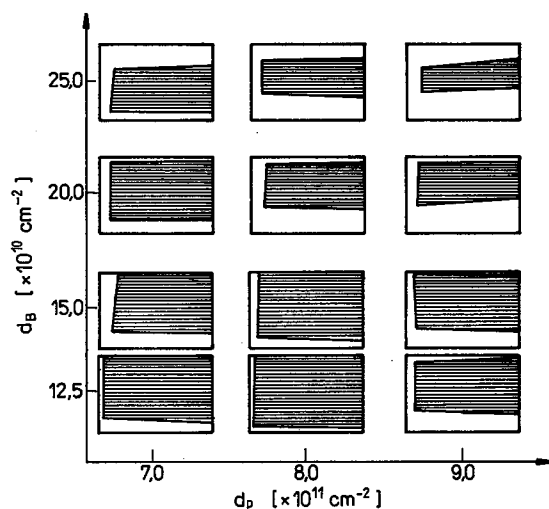
Dla wyboru parametrów krytycznych do badań obszarów funkcjonalnych zastosowano zbliżoną do opisaną [11] procedurę doświadczalnego badania czułości mierzonych obszarów dla różnych wartości podstawowych parametrów technologicznych. Na podstawie analizy wyników przeprowadzonych badań wybrano dwa



Rys. 19. Obszar funkcjonalny struktury drugiego układu wytworzonej zgodnie z parametrami początkowego punktu technologicznego



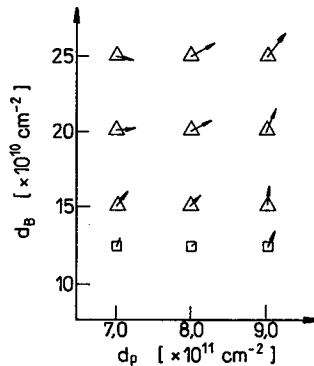
Rys. 20. Ilustracja współrzędnych dwóch planów eksperymentów dla układu drugiego



Rys. 21. Mapa obszarów funkcjonalnych badanych układów wytworzonych w punktach odpowiadających planowi pierwszego i drugiego eksperymentu

parametry dynamiczne, które wraz z napięciem zasilającym stanowią parametry krytyczne opracowywanego układu. Parametry te określają odpowiednio: czas opóźnienia danych względem sygnału sterującego ich zapisem do układu i okres zegara zewnętrznego  $t_{CY}$  synchronizującego pracę całego układu (pokazano tylko obszary funkcjonalne dla parametru  $t_{CY}$ ). Biorąc pod uwagę podobieństwa technologiczne, różnice konstrukcyjne oraz zmierzone obszary funkcjonalne dla struktur w punkcie startowym (rys. 19), zaproponowano plan eksperymentu czynnikowego kompletnego, którego schemat w przestrzeni optymalizowanych parametrów pokazano na rys. 20 (punkty  $\Delta$ ).

Po dokonaniu pomiarów obszarów funkcjonalnych wyznaczono wektory kierunkowe dla poszczególnych punktów technologicznych. Na rysunku 21 pokazano zbiorcze obszary funkcjonalne (dla uproszczenia względem jednego z parametrów  $t_{CY}$ ), zaś rys. 22 wyznaczone wektory kierunkowe w poszczególnych punktach planu. Analizu-



Rys. 22. Mapa wektorów kierunkowych dla punktów z dwóch planów eksperymentów dla układu drugiego

jąc przedstawione obszary można stwierdzić, że w żadnym punkcie nie jest spełniony warunek całkowitego pokrycia okna pomiarowego przez badany obszar. Wobec tego zdecydowano się przeprowadzić uzupełniający eksperyment technologiczny, którego plan zaznaczono schematycznie na rys. 20 (punkty  $\square$ ). Po wytworzeniu partii płytek z parametrami określonymi przez eksperyment uzupełniający i dokonaniu typowych pomiarów obszarów funkcjonalnych (rys. 21) wyznaczono wektory kierunkowe, których wielkości zaznaczono na rys. 22. Na podstawie analizy obszarów funkcjonalnych i wektorów kierunkowych dla obu planów stwierdzono, że maksymalne pokrycie okna pomiarowego występuje w punkcie planu o współrzędnych:  $d_B = 1.25 \times 10^{11} [1/\text{cm}^2]$  i  $d_p = 8.0 \times 10^{11} [1/\text{cm}^2]$ .

## ZAKOŃCZENIE

Celowość opracowania przedstawionej metody centrowania technologicznego wyniknęła z analizy przyczyn stosunkowo niskiego uzysku produkcyjnego w odniesieniu do układów wielkiej skali integracji. Wśród wielu różnych przyczyn tego faktu,

na szczególną uwagę zasługują te, które leżą w sferze projektowania układu bowiem przy niewłaściwym projekcie wszelkie zabiegi technologiczne zmierzające do zwiększenia uzysku są nieefektywne, gdyż nie można przekroczyć pewnej wartości granicznej uzysku wynikającej z konkretnej realizacji funkcji układowej.

Prezentowana metoda jest heurystyczną metodą doświadczalną. Opracowanie metody tego typu, w dobie powszechnej komputeryzacji może w pierwszej chwili budzić wątpliwości, jednakże po szczegółowej analizie dostępnych programów modelowania i symulacji zarówno z punktu widzenia dokładności jak i dopuszczalnej złożoności obliczanych układów scalonych, okazuje się że badania eksperymentalne mogą z powodzeniem stanowić uzupełnienie projektowania symulacyjnego. Powyższy wniosek jest tym bardziej uzasadniony, jeśli weźmie się pod uwagę aspekty ekonomiczne obu typów projektowania. Porównanie czasu i kosztów wykonania serii eksperymentów z czasem i kosztami kilkutygodniowej pracy dużego komputera wypada na korzyść metod eksperymentalnych.

Śluszność i skuteczność proponowanej metody potwierdziły jej praktyczne realizacje w trakcie opracowywania układów scalonych. W przypadku układów opracowanych z wykorzystaniem metody centrowania technologicznego, w znacznie krótszym czasie osiąga się poziom uzysku produkcyjnego przewyższający poziom uzysku układów projektowanych tradycyjnie. Pierwotnie metoda centrowania technologicznego została opracowana dla układów cyfrowych wielkiej skali integracji, niemniej jednak może być z powodzeniem stosowana również do układów analogowych lub analogowo-cyfrowych projektowanych w różnych technologiach.

#### BIBLIOGRAFIA

1. W. Ham: *Yield area analysis: Part I — A diagnostic tool for fundamental integrated circuit process problems*. RCA Rev., vol. 39, no. 2, pp. 231—249, 1978
2. W. Mały, A. Strójas: *Statistical simulation of the IC manufacturing process*. IEE Trans. on CAD of Integr. Cir. and Sys., vol. CAD-1, no. 3, pp. 120—131, 1982
3. S. Director, L. Vidigal: *Statistical circuit design: A somewhat biased survey*. Proc. of the 1981 ECCTD Conf., pp. 15—24, Hague. The Netherlands, 1981
4. R. Brayton at all: *A survey of optimization techniques for integrated circuit design*. Proc. of the IEEE, vol. 69, no. 10, pp. 1334—1362, 1981
5. T. Yu at all: *Statistical performance modeling and parametric yield estimation of MOS VLSI*. IEEE Trans. on CAD of Integr. Cir. and Sys., vol. CAD-6, no. 6, pp. 1013—1022, 1987
6. S. Nassif at all: *FABRICS II. A statistically based IC fabrication process simulator*. IEEE Trans. on CAD of Integr. Cir. and Sys., vol. CAD-3, no. 1. pp. 40—46, 1984
7. *Customers calm as Intel gets 386 back on track*. Electronics, vol. 60, no. 12, pp. 35—36, 1987
8. Z. Polański: *Planowanie doświadczeń w technice*. Warszawa: PWN, 1984
9. Y. Aoki at all: *New design centering methodology for VLSI device development*. IEEE Trans. on CAD of Integr. Cir. and Sys., vol. CAD-6, no. 3, pp. 452—461, 1987
10. M. Stybliński: *Metody analizy i optymalizacji statystycznej układów elektronicznych*. Prace Naukowe PW, z. 53, Warszawa, 1981
11. M. Patyra: *Metoda optymalizacji parametrów technologicznych monolitycznych układów scalonych*. Rozpr. Dokt., Wyd. Elektroniki PW, Warszawa, 1985

12. B. Kamińska, J. Zabrodzki: *The concept of proper operation region in digital integrated circuit testing, the 8080A microprocessor as an example*. Microproc. and Microprogr., vol. 12, no. 2, pp. 285—290, 1983

MAREK J. PATYRA, JAN ZABRODZKI

TECHNOLOGICAL PARAMETERS CENTERING OF  
LSI AND VLSI INTEGRATED CIRCUITS.  
EXPERIMENTAL APPROACH

S u m m a r y

One of the basic demand of the design process integrated circuits is to ensure the proper production yield. The tuning of technological parameters is described with the real condition of a given line taken into account. This problem is called technological centering. The experimental-heuristic method of searching the solution of IC technological centering task is proposed. The method enable practical verification of IC project obtained by using analithical-simulation procedure and optimization of this project in the real technological line. The practical results of proposed the method application are presented. The results shown in the paper confirm the purpose of application of technological centering in practice.

# Rekonstrukcja obrazów adaptacyjnymi filtrami liniowymi

JAN RADECKI, TOMASZ CHADY

*Instytut Elektroniki i Informatyki, Politechnika Szczecińska*

*Otrzymano 1988.12.31*

*Autoryzowano do druku 1989.06.30*

W artykule przedyskutowano jednowymiarowy adaptacyjny filtr rekursywny w zastosowaniu do poprawiania stosunku sygnał/szum w obrazach zakłóconych szumem o rozkładzie Gaussa. Załączono przykłady filtracji obrazów zakłóconych. Przetwarzaniu poddano obrazy o małym współczynniku sygnał/szum.

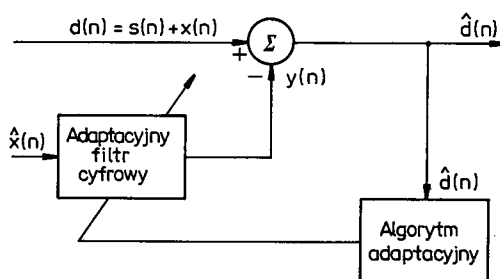
## 1. WSTĘP

Podstawowe problemy w rekonstrukcji obrazów, tj. poprawienie stosunku sygnał/szum oraz zmniejszenie efektu odpowiedzi impulsowej liniowego systemu, są rozwiązywane głównie w oparciu o techniki adaptacyjne. Metody adaptacyjnego przetwarzania rozkładów przestrzennych są uwarunkowane wieloma czynnikami takimi jak: charakter zakłóceń, przyjęty model obrazu [6], dostępność sygnału odniesienia lub jej brak, wymagania odnośnie szybkości przetwarzania systemu itd. Do obniżenia poziomu szumu przy jednoczesnym zachowaniu krawędzi w obrazie są stosowane filtry liniowe [1], [2], [3] oraz filtry nieliniowe [7], a głównym kryterium wyboru struktury są parametry stochastyczne sygnału zakłócającego. Stosunkowo dobre rezultaty w rekonstrukcji obrazu zakłóconego szumem o rozkładzie Gaussa uzyskano stosując filtry Kalmana dwuwymiarowe [1], [2], a także kaskadę filtrów jednowymiarowych rotowanych o odpowiednie kąty [1], [3]. Efektywną metodą eliminacji zakłóceń impulsowych o dużej amplitudzie i małym prawdopodobieństwie wystąpienia, jest zastosowanie 2-D filtrów nieliniowych. W wielu zastosowaniach istnieje możliwość obserwacji zakłócenia i na podstawie znajomości jego parametrów statystycznych dokonuje się estymacji sygnału wolnego od szumu. Metoda ta znalazła szerokie zastosowanie w układach 1-D głównie przy zastosowaniu filtrów nie-rekursywnych [4], [5], [8].

W pracy rozważono zagadnienie rekonstrukcji obrazu zakłóconego szumem o rozkładzie Gaussa przy pomocy liniowych filtrów adaptacyjnych o strukturze rekursywnej. Załączono przykład rekonstrukcji obrazów.

## 2. ELIMINACJA SZUMU Z ROZKŁADU PRZESTRZENNEGO — JEDNOWYMIAROWY FILTR REKURSYWNY

Dla poprawienia stosunku sygnał/szum w obrazie zakłóconym szumem o rozkładzie Gaussa, tj. szumem o wartości oczekiwanej  $E[X(m, n)] = 0$  i stałej wariancji, zastosowano jednowymiarowy (1—D) filtr rekursywny. Przyjęto tu wariant z dostępnym szumem odniesienia [3], [4]. Zastosowany filtr rekursywny daje większą szybkość przetwarzania i krótszy czas adaptacji w porównaniu z filtrem nierekursywnym, ale problemy stabilności i nieliniowości fazy stwarzają trudności aplikacyjne tego modelu. Schemat układu adaptacyjnego przedstawiono na rys. 1. Filtr adapta-



Rys. 1. Schemat filtru adaptacyjnego

cyjny przetwarza sygnał odniesienia  $\hat{x}(n)$  skorelowany w nieznany sposób z szumem  $x(n)$  zakłócającym addytywnie sygnał  $s(n)$ . Rozkłady przestrzenne  $X(m, n)$  i  $D(m, n)$  przekształcono w jednowymiarowe ciągi poprzez wybieranie próbek linia po linii wg schematu:

$$d_{(n)} = [d_{(11)}, \dots, d_{(n)}, \dots, d_{(1N)}; \\ d_{(m1)}, \dots, d_{(mn)}, \dots, d_{(mN)}; \\ d_{(N1)}, \dots, d_{(NN)}, \dots, d_{(NN)}] \quad (1)$$

Należy oczywiście wziąć pod uwagę utratę pewnych przestrzennych powiązań przy jednowymiarowej analizie obrazu.

Sygnał odniesienia  $\hat{x}(n)$  jest przetworzony przez filtr adaptacyjny tak aby sygnał na wyjściu układu  $\hat{d}_{(n)} = d_{(n)} - y_{(n)}$  był najlepszą estymatą sygnału oryginalnego  $s_{(n)}$  w sensie minimum błędu średniokwadratowego. Przy czym  $d_{(n)}$  jest zakłóconą wersją sygnału  $s_{(n)}$ . Powyższy warunek jest spełniony jeżeli moc wyjściowa układu jest minimalna [9]. Minimalizację mocy wyjściowej uzyskuje się minimalizując wartość oczekiwaną kwadratu sygnału na wyjściu układu  $\hat{d}_{(n)}$ , będącego różnicą między sygnałem dostępnym (mierzonym)  $d_{(n)}$ , a sygnałem  $y_{(n)}$  na wyjściu filtra:

$$\hat{d}_{(n)} = d_{(n)} - y_{(n)}, \quad (2)$$

gdzie:

$$y(n) = \sum_{k=0}^p a_k x_{(n-k)} - \sum_{k=1}^q b_k y_{(n-k)} = A^T X - B^T Y \quad (3)$$

stąd:

$$\begin{aligned} E[\hat{d}_{(n)}^2] &= \sum_{l=0}^L \hat{d}_{(n-L)}^2 = \varepsilon = \\ &= E[d_{(n)}^2] + A^T R_{xx} A + B^T R_{yy} B - 2A^T R_{dx} + 2B^T R_{dy} - 2A^T R_{xy} B \end{aligned} \quad (4)$$

gdzie:

$R$  — operator korelacji poszczególnych sygnałów.

$E$  — operator wartości oczekiwanej.

Przy założeniu stacjonarności sekwencji  $\{d_{(n)}\}$  i  $\{x_{(n)}\}$  oraz niezależności sygnałów  $x_{(n)}$  i  $d_{(n)}$ , wektor odpowiedzi impulsowej filtru Wienera jest zdefiniowany:

$$H_{xy} = R_{xx}^{-1} R_{xy}, \quad (5)$$

gdzie:

$R_{xx} = E[X_{(n)} X_{(n)}^T]$  — macierz autokorelacji sygnału wejściowego.

$R_{xy} = E[Y_{(n)} X_{(n)}^T]$  — macierz interkorelacji sygnału wejściowego i wyjściowego filtru.

Zakłada się, że nieujemnie określona macierz  $R$  jest odwracalna. Wartości współczynników modelu uaktualniano metodą minimalizacji gradientowej tj.:

$$a^{j+1} = a^j - \mu_A \nabla A \quad (6)$$

$$b^{j+1} = b^j - \mu_B \nabla B, \quad (7)$$

przy czym:

$$\nabla A = \frac{\partial E[\hat{d}_{(n)}^2]}{\partial A}; \quad \nabla B = \frac{\partial E[\hat{d}_{(n)}^2]}{\partial B}. \quad (8)$$

Biorąc pod uwagę, że korelację np.

$$R_{xx} = E[X_{(n)} X_{(n)}^T] \quad (9)$$

sygnału ergodycznego można wyrazić wzorem:

$$R = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N X_{(n)} X_{(n)}^T \quad (10)$$

i korzystając z wzorów: (6), (7), (9), (10), otrzymujemy po przekształceniach:

$$a_k^{j+1} = a_k^j + 2\mu_A \frac{1}{K} \sum_{k=0}^K \hat{d}_{(n-k)} \hat{x}_{(n-k)}, \quad (11)$$

$$b_k^{j+1} = b_k^j - 2\mu_B \frac{1}{K} \sum_{k=0}^K \hat{d}_{(n-k)} y_{(n-k)}. \quad (12)$$

Tak sformułowany algorytm jest wolny od wpływu sygnału użytecznego na wartość gradientu zmian współczynników gdyż:

$$\sum_{k=0}^{K-1} \hat{d}_{(n-k)} \hat{x}_{(n)} = \sum_{k=0}^{K-1} (s_{(n-k)} + x_{(n-k)} - y_{(n-k)}) \hat{x}_{(n)} \quad (13)$$

sygnał  $s_{(n)}$  nie jest skorelowany z  $\hat{x}_{(n)}$

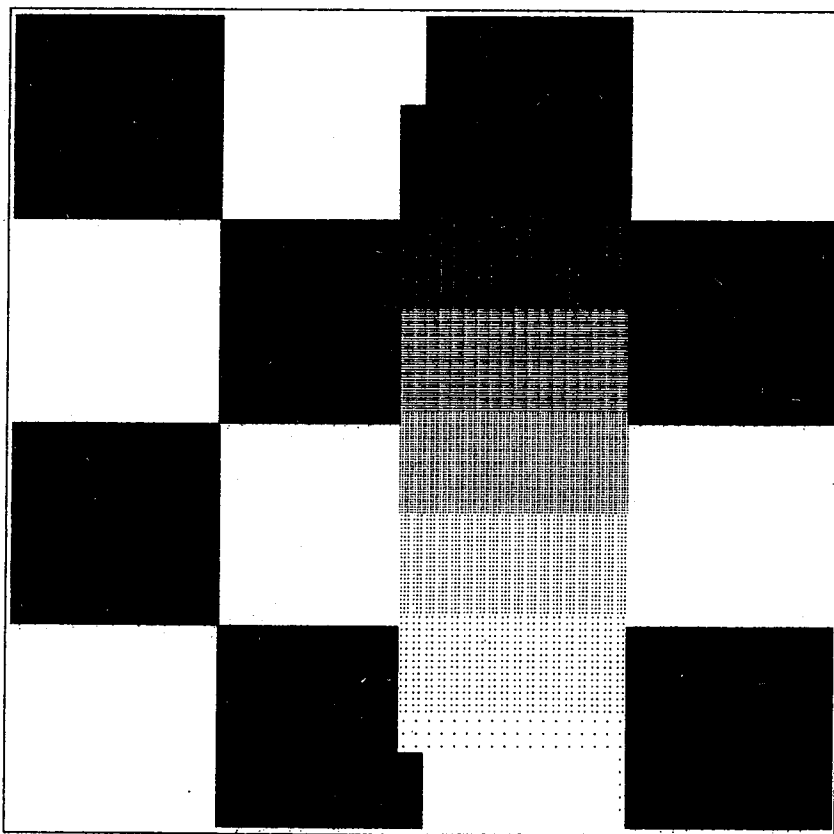
$$\sum_{n=0}^{K-1} s_{(n-k)} \hat{x}_{(n)} = 0 \quad (14)$$

i stąd

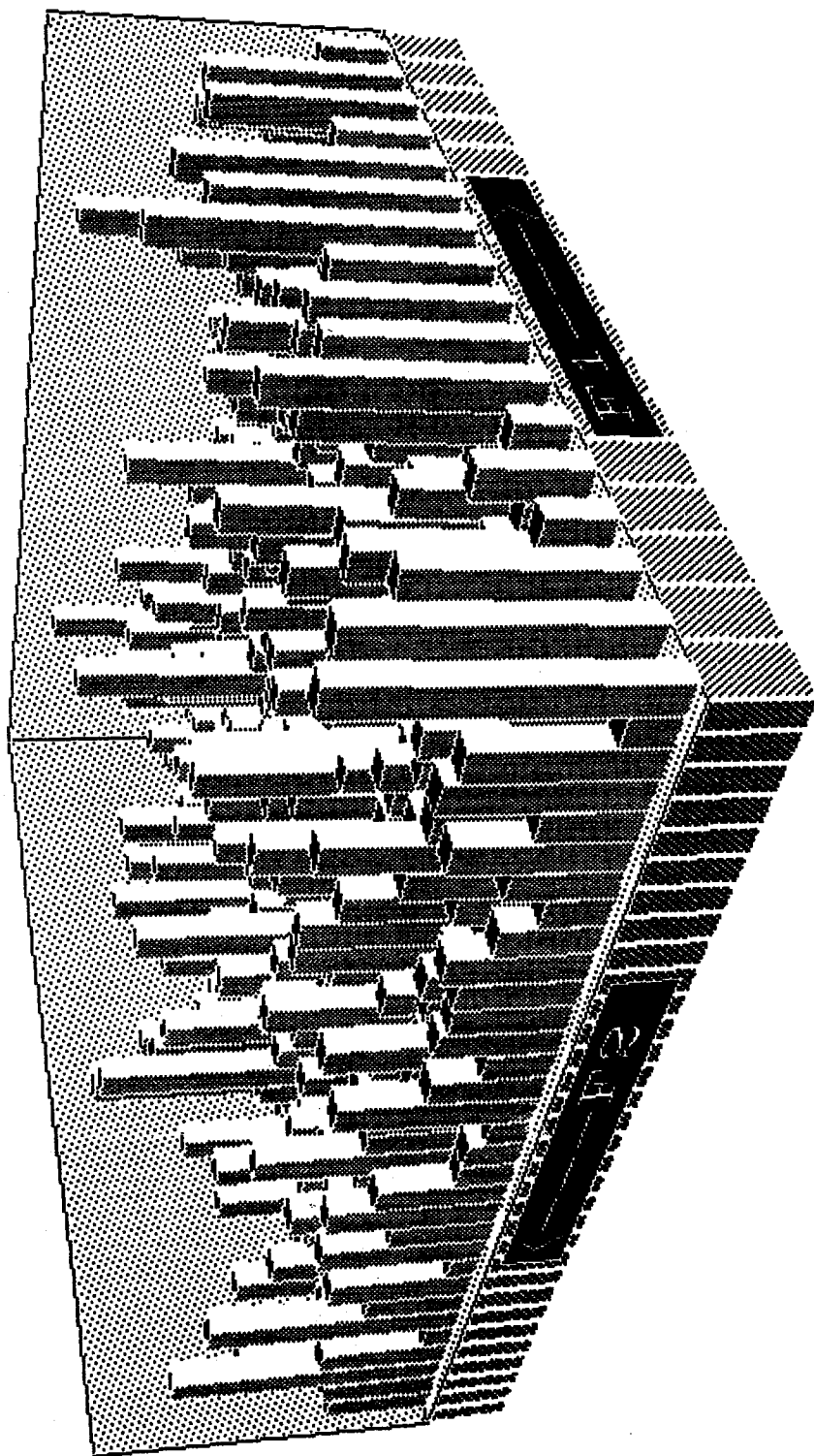
$$\sum_{n=0}^{K-1} \hat{d}_{(n-k)} x_{(n)} = \sum_{n=0}^{K-1} (x_{(n-k)} - y_{(n-k)}) \hat{x}_{(n)}, \quad (15)$$

czyli jest niezależny od sygnału użytecznego.

$\mu$  jest parametrem ustalającym szybkość zbieżności algorytmu oraz decyduje o wartości błędu średniokwadratowego adaptacyjnego filtra. W stanie ustalonym elementy wektora  $H$  oscylują wokół rozwiązania optymalnego  $H_{opt}$  co jest przyczyną błędu średniokwadratowego i błąd ten jest proporcjonalny do kroku adaptacji  $\mu$ . Wybór wartości  $\mu$  jest zatem kompromisem między szybkością zbieżności algorytmu oraz wartością błędu średniokwadratowego. Czas adaptacji jest również proporcjonalny do ilości elementów wektora  $H$ . Duża liczba elementów wektora  $H$  i mała wartość  $\mu$  dają mały błąd średniokwadratowy między sygnałem użytecznym  $s_{(n)}$  a jego estematą  $\hat{d}_{(n)}$  ale czas zbieżności algorytmu jest długi.



Rys. 2. Obraz testowy



Rys. 3. Widmo szumu

[143]

Warunkiem koniecznym dla zbieżności algorytmu jest aby:

$$0 \leq \mu \leq \frac{1}{\lambda_{\max}} \quad (16)$$

gdzie:

$\lambda_{\max}$  — największa wartość własna autokorelacji  $R_{xx}$ .

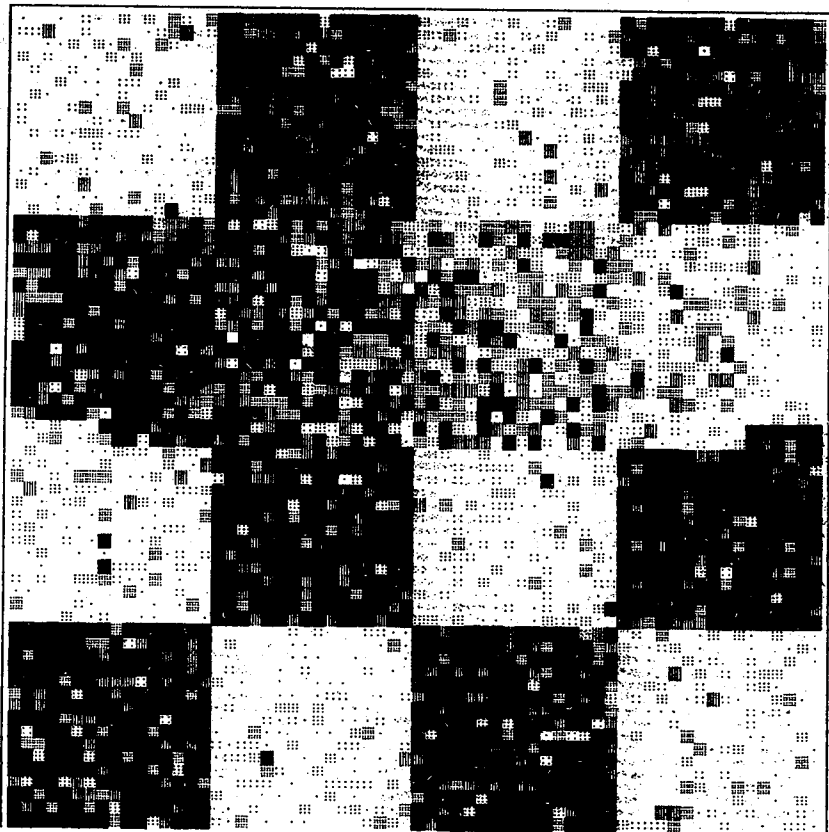
Powyższy model zastosowano do eliminacji szumu z obrazu w warunkach małej wartości współczynnika sygnał/szum. Badania przeprowadzone na obrazie testowym o ośmiu stopniach gradacji i wymiarach  $M \times N = 64 \times 64$  punktów (rys. 2). Obraz zakłócono szumem 2-D o rozkładzie Gaussa, którego widmo zamieszczono na rys. 3., przy czym stosunek sygnał/szum w obrazie zakłóconym (rys. 4.) wynosi 6 dB.

Filtrację obrazu zakłóconego dokonano filtrem 1-D opisanym równaniem (3) przy czym:

$$A^T = [a_0, a_1, a_2]; B^T = [1, b_1, b_2] \quad (17)$$

Wartości startowe współczynników były równe współczynnikom transmitancji aproksymowanej funkcjami Bessela przy  $f_c/f_s = 0.3$

gdzie:



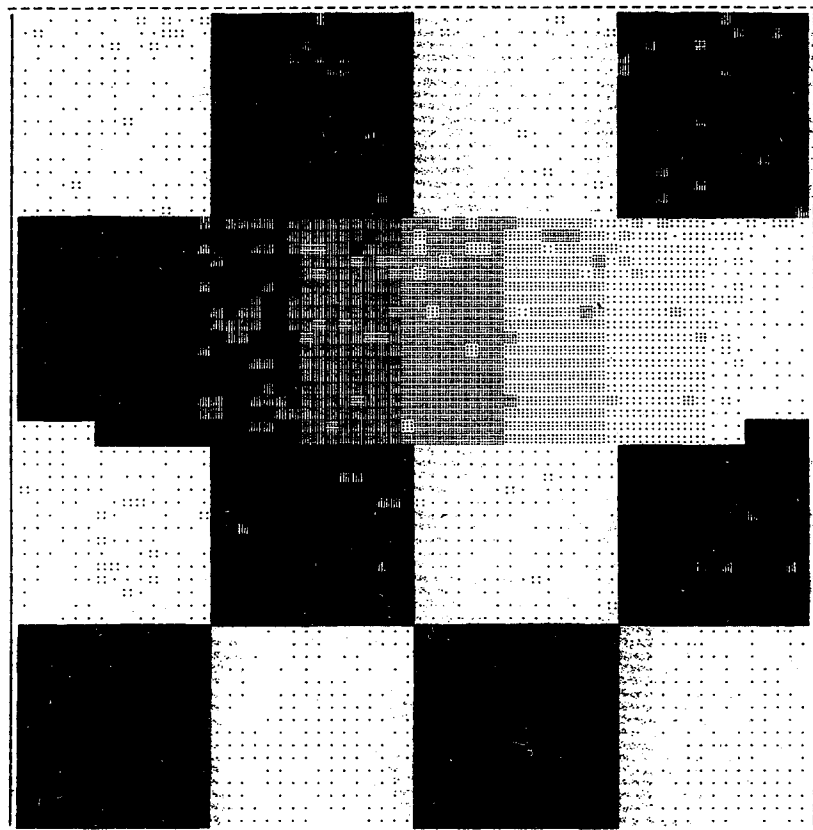
Rys. 4. Obraz zakłócony szumem o widmie przedstawionym na rys. 3.

$f_c$  — częstotliwość graniczna filtru.

$f_s$  — częstotliwość próbkowania.

Przy opracowywaniu algorytmów adaptacyjnego cyfrowego przetwarzania obrazu, zwykle jest przyjmowane założenie, że obraz jest rozkładem stacjonarnym. Przyjęcie takiego modelu powoduje, iż występują duże różnice między obrazem rzeczywistym, a modelem ponieważ większość obrazów jest z natury niestacjonarna. Przyjmuje się, że obraz jest globalnie stacjonarny z lokalnymi niestacjonarnościami i algorytm adaptacyjny winien nadążać za lokalnymi zmianami parametrów statystycznych przetwarzanego sygnału. Wartość współczynnika kroku  $\mu$  (16) wpływa na szybkość nadążania za lokalnymi zmianami parametrów statystycznych obrazu i stąd właściwy wybór  $\mu$  jest kompromisem między szybkością zbieżności algorytmu, nadążaniem za zmianami statystycznymi w obrazie oraz błędem adaptacji mierzonym jako oscylacje wokół optymalnego rozwiązania. Wzrost wartości  $\mu$  powoduje wzrost szybkości zbieżności algorytmu lepszą adaptację w lokalnych obszarach niestacjonarnych, ale rośnie błąd adaptacji. W rozważanym przypadku wybrano eksperymentalnie wartości:

$$\mu_A = 0.02; \mu_B = 0.005$$



Rys. 5. Obraz z rys. 4 przetworzony filtrem adaptacyjnym

Takie wartości współczynników kroku dają względnie dużą szybkość zbieżności algorytmu (6), (7) i spełniają warunek (16). Małe zmiany w kolejnych iteracjach współczynników mianownika transmitancji w zasadzie zapewniają stabilność filtru. Stosunek sygnał/szum obliczany według wzoru:

$$SNR = 10 \log \frac{\delta^2}{e} [\text{db}] \quad (18)$$

$$e = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (S_{i,j} - D_{i,j})^2, \quad (19)$$

gdzie:

$\delta^2$  jest wariancją oryginalnego sygnału (bez zakłóceń);

$\{S_{i,j}\}$  i  $\{D_{i,j}\}$  reprezentują odpowiednio  $N \times N$  oryginalny i przetworzony obraz. Sygnał wynosił w obrazie przed filtracją  $SNR = 6$  dB, a po filtracji  $SNR = 21$  dB. Obraz po filtracji przedstawiono na rys. 5.

Przedstawiona metoda przetwarzania zakłóconych rozkładów przestrzennych pozwala na uzyskanie znacznej poprawy jakości obrazu pod warunkiem że przetwarzane sygnały są globalnie stacjonarne. Jednowymiarowy model filtru umożliwia zastosowanie struktury rekursywnej co zdecydowanie skraca czas adaptacji i przetwarzania filtru.

## BIBLIOGRAFIA

1. J. W. Woods, V. K. Ingle: *Kalman Filtering in Two Dimensions: Further Results*. IEEE Trans. Acoust., Speech, Signal. Proc., vol. ASSP-29, No. 2. April 1981
2. J. W. Woods, S. Dravida, R. Mediavilla: *Image Estimation Using Doubly Stochastic Gaussian Random Field Models*. IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-9, No. 2. March 1987
3. P. Chan, J. S. Lim: *One-Dimensional Processing for Adaptive Image Restoration*. IEEE Trans. Acoust., Speech, Signal. Proc., vol. ASSP-33, No. 1. Feb. 1985
4. N. J. Bershad, P. L. Feintuch: *The Recursive Adaptive LMS Filter—A Line Enhancer. Application and Analytical Model for the Mean Weight Behaviour*. IEEE Trans. Acoust., Speech, Signal. Proc., vol. ASSP-32, No. 1. Feb. 1984
5. P. L. Feintuch: *An Adaptive Recursive LMS Filters*. Proc. IEEE, vol. 64, Nov. 1976
6. S. Geman, D. Geman: *Stochastic Relaxation, Gibbs' Distribution, and the Bayesian Restoration of Images*. IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, No. 6. Nov. 1984
7. J. Radecki, E. Kornatowski, T. Chady, H. Kubiak: *Technika Adaptacyjna w Przetwarzaniu Sygnałów Dwuwymiarowych*. X Krajowa Konferencja „Teoria Obwodów i Układy Elektroniczne”, Gdańsk 21—24 X, 1987
8. T. A. C. M. Claasen, W. F. G. Mecklenbrauker: *Adaptive Techniques for Signal Processing in Communications*. IEEE Communication Magazine, vol. 23, No. 11. 1985
9. B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. M. Hearn, J. R. Zeidler: *Adaptive noise cancelling; Principles and Applications*. Proc. IEEE, vol. 63, pp. 1692—1716, Dec. 1975

J. RADECKI, T. CHADY

## ADAPTIVE LINEAR FILTER IN IMAGE RESTORATION

## Summary

The paper presents a monodimensional approach for adaptive processing of images disturbed by Gaussian noise. An iterative frequency domain technique is presented for estimating autoregressive moving average parameters. The experimental result based on a simulated image demonstrating the filtering effectiveness of the adaptive algorithm is included.



# Elektryczne metody badania aerozoli

ANATOL JAWOREK

*Instytut Maszyn Przepływowych PAN, Gdańsk*

*Otrzymano 1988.11.15*

*Autoryzowano do druku 1989.06.03*

Przedstawiono przegląd podstawowych metod elektrycznych stosowanych w badaniach wielkości cząstek stałych lub kropeł cieczy tworzących aerosol. Opisano układ kondensatora płaskiego Straubela oraz elektrostatyczne metody sedymentacyjne służące do pomiaru wielkości pojedynczych cząstek. Przedstawiono też metody separacji elektrostatycznej umożliwiające wyznaczenie gęstości lub dystrybuanty rozkładu wielkości cząstek. Z pozostałych metod elektrycznych opisano sondy drutowe i igłowe oraz piezoelektryczne.

## PRZEDMOWA

W pracy przedstawiono przegląd podstawowych, elektrycznych metod pomiarowych stosowanych do badania wielkości cząstek stałych lub kropeł cieczy zawieszonych w gazie. Opisano zarówno metody pozwalające uzyskać sygnał elektryczny jak i te, które posługują się zjawiskami elektrycznymi do separacji cząstek. Zasady leżące u podstaw metod elektrycznych oraz metod optycznych wykorzystywane są coraz powszechniej w budowie automatycznych urządzeń pomiarowych.

Od czasów doświadczenia wykonanego przez Millikana (które wprawdzie nie służyło do badania aerozoli, ale do tego celu zostało przez późniejszych badaczy zaadaptowane) opracowano wiele przyrządów i metod opartych na zasadzie elektryzowania aerozolu i badania zachowania się poszczególnych tworzących go cząstek poruszających się w polu elektrycznym lub będących w spoczynku (w stanie zawieszenia).

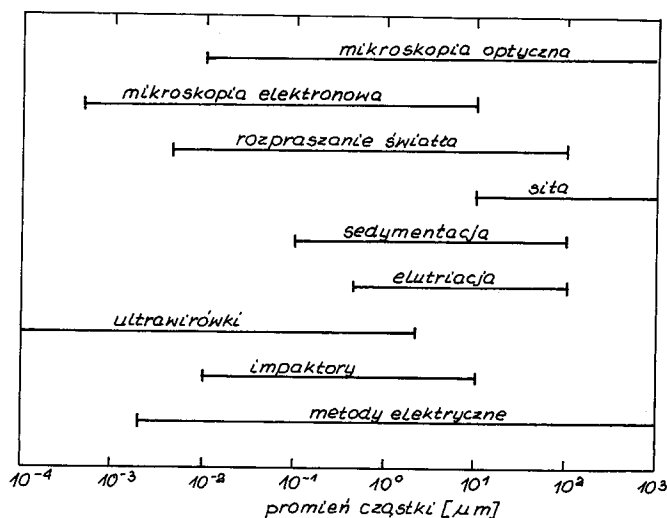
Pomimo tak dużego znaczenia jakie zyskują te, bardzo często już klasyczne metody (o czym świadczą katalogi firm produkujących aparaturę badawczą np. [1, 2], brak jest w literaturze pracy systematyzującej liczne pomysły rozwiązań w tej dziedzinie.

Podstawowy podręcznik w literaturze polskiej [3] traktujący o rozpylaniu cieczy, z elektrycznych metod badawczych wspomina tylko o sondzie drutowej i igłowej. Autorzy artykułów przeglądowych, poświęconych metodom analizy wielkości cząstek [4, 5, 6, 7, 8] koncentrują się na metodach hydrodynamicznych (impaktory kaskadowe, wirówki odśrodkowe), metodom elektrycznym poświęcając znacznie mniej miejsca lub tylko wzmiankując ich istnienie.

Praca ta jest próbą uzupełnienia tej luki i zebrania podstawowych idei leżących u podstaw współczesnych elektrycznych metod badawczych, rozproszonych w różnych publikacjach, chociaż zawarty w niej materiał na pewno nie jest kompletny.

## 1. WSTĘP

Poznanie rozkładu wielkości cząsteczek aerozolu ma istotne znaczenie w wielu dziedzinach badawczych i procesach przemysłowych. Zagadnienie prostego i szybkiego pomiaru parametrów zawieszin cząstek stałych i cieczy, w tym cząstek naelektryzowanych nie zostało jednak dotychczas należycie rozwiązane. Wśród licznych metod badania wielkości cząstek, pokrywających różne zakresy [4—8] (rys. 1), szczególnie



1. Zakres zastosowań podstawowych metod badania aerozoli

miejsce przypada metodom elektrycznym, które obok metod optycznych lub w połączeniu z nimi umożliwiają uzyskanie sygnału elektrycznego nadającego się do bezpośredniej rejestracji lub przetwarzania. Pomimo niewątpliwych zalet obu metod i rozwiniętej produkcji urządzeń służących do tego celu [1, 2] żadne z rozwiązań nie ma znaczenia uniwersalnego pozwalającego pokryć cały zakres możliwych wielkości i koncentracji cząstek oraz nie daje się stosować do wszystkich rodzajów cząstek. Dlatego opracowano wiele różnorodnych metod i technik badawczych.

Najbardziej uniwersalne są metody optyczne [9], ale można je stosować tylko w przypadku rozróżnialności poszczególnych cząstek aerozolu. Gdy koncentracja cząstek przekroczy pewną wartość graniczną, możliwy jest tylko pomiar wartości uśrednionych. W przypadku cząstek naładowanych, metodami optycznymi można określić tylko wielkość cząstek. Gdy zachodzi potrzeba dodatkowego określenia ładunku niesionego przez cząstki lub ich ruchliwości, można stosować metody elektryczne, zaś metody optyczne wymagają wspomagania elektrycznego.

Jako pierwszą przedstawiono metodę Straubela, będącą zmodyfikowanym doświadczeniem Millikana, w którym możliwy był pomiar tylko ładunku właściwego cząstek. Kondensator Straubela umożliwia niezależne określenie ładunku i wielkości cząstek. Najwięcej miejsca poświęcono metodom elektrostatycznym, do których zaliczają się metoda sedymentacyjna i separacji elektrostatycznej. W metodzie sedymentacyjnej obserwuje się (fotografuje) tory naładowanych cząstek opadających w poprzecznym polu elektrycznym stałym lub zmiennym. Na tej podstawie można określić wielkość i ładunek cząstek. W metodzie separacji elektrostatycznej badane cząstki rozdziela się przestrzennie w zależności od niesionego przez nie ładunku, który jest proporcjonalny do powierzchni cząstek. Z pozostałych metod elektrycznych opisano sondę drutową i igłową. W sondzie drutowej obojętny elektrycznie aerozol może odbierać ładunek z drutu, który znajduje się na wysokim potencjale lub oddawać zgromadzony na swej powierzchni ładunek elektryczny (cząstki są uprzednio elektryzowane). Wielkość wytworzonego w obu przypadkach impulsu jest miarą wielkości cząstki. W sondzie igłowej przelatująca kropla cieczy zwiera dwa ustawione naprzeciw siebie końce igieł powodując przepływ prądu w obwodzie elektrycznym. Oba typy sond przydatne są tylko do badania cząsteczek przewodzących. Na zakończenie przedstawiono czujniki piezoelektryczne, które umożliwiają określenie wielkości cząstki na podstawie wielkości impulsu wytworzonego przez cząstkę przy jej zderzeniu z sondą lub na podstawie zmiany częstotliwości drgań własnych piezoelektryka przy zmianie jego masy przez osiadającą cząstkę.

## 2. PODSTAWOWE ZALEŻNOŚCI OPISUJĄCE RUCH CZĄSTEK

Wiele z przedstawionych metod opiera się na ruchu naelektryzowanych cząstek w polu elektrostatycznym. Pole to może być prostopadłe lub równoległe do sił grawitacji. W wielu przypadkach wprowadza się dodatkowo strumień powietrza, który unosi badane cząstki.

Równania ruchu cząstek w stanie ustalonym wyznacza się na podstawie równowagi sił działających na cząstkę:

— siły grawitacji

$$F_g = \frac{4}{3} \pi r^3 \rho_p g, \quad (1)$$

— siły Stokesa

$$F_s = 6\pi\eta r(v_a - v_p), \quad (2)$$

— siły wywieranej przez pole elektrostatyczne

$$F_e = QE \quad (3)$$

przy czym:  $r$  — promień cząstki,  $Q$  — ładunek zebrany na cząstce,  $\rho_p$  — gęstość cząstki,  $g$  — przyspieszenie ziemskie,  $\eta$  — lepkość gazu (powietrza),  $v_a$ ,  $v_p$  — prędkość powietrza i cząstki,  $E$  — natężenie pola elektrycznego. W równaniach tych oraz w dalszych rozważaniach założono kulisty kształt cząstek.

Zależności (1) — (3) pozwalają wyznaczyć prędkość ruchu cząstek w obecności pola elektrostatycznego a następnie wielkość badanych cząstek i ich ładunek. Na przykład prędkość cząstki w nieruchomym powietrzu ( $v_a = 0$ ) w kierunku pola elektrostatycznego wyznaczona z (2) i (3) wynosi:

$$v_p = \frac{QE}{6\pi\eta r}. \quad (4)$$

Prędkość sedymentacji cząstek pod wpływem sił grawitacji wyznaczona z (1) i (2) wynosi:

$$v_p = \frac{2r^2\varrho_p g}{9\eta}. \quad (5)$$

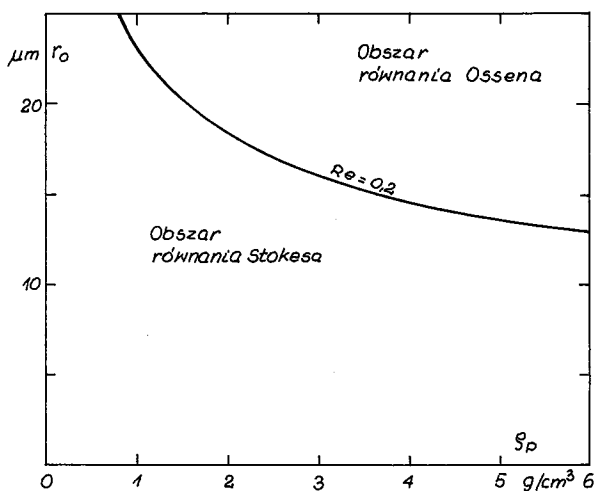
Chociaż równania te spełnione są tylko w warunkach idealnych, mogą być one używane do opisu większości praktycznych przypadków. W analizie niektórych urządzeń pomiarowych muszą być jednak uwzględnione dodatkowe czynniki.

Dla cząstek o średnicy porównywalnej ze średnią drogą swobodną molekuł gazu  $\lambda$  tzn. dla liczb Knudsena bliskich lub mniejszych od jednośc, należy wprowadzić czynnik korekcyjny Cunninghama [7]:

$$v_p = \frac{2(\varrho_p - \varrho_a)gr^2}{9\eta} \left( 1 + \frac{A\lambda}{2r} \right) \quad (6)$$

gdzie:  $A = 0,85$ . Czynnik ten uwzględnia swobodny ruch cząstki aerozolu pomiędzy kolejnymi zderzeniami z molekułami gazu.

Równanie (5) jest spełnione dla liczb Reynoldsa mniejszych od 0,2. Dla większych liczb Reynoldsa stosuje się zależność Oseena [7]:



2. Zakresy stosowania równań Stokesa i Oseena

$$v_p = \frac{1}{2} \left( \frac{K_1}{r} + \sqrt{\left( \frac{K_1}{r} \right)^2 + 4rK_2} \right), \quad (7)$$

w której

$$K_1 = \frac{8\eta}{3\varrho_a}, \quad K_2 = \frac{16}{27} \frac{\varrho_a - \varrho_p}{\varrho_a} g \quad (8)$$

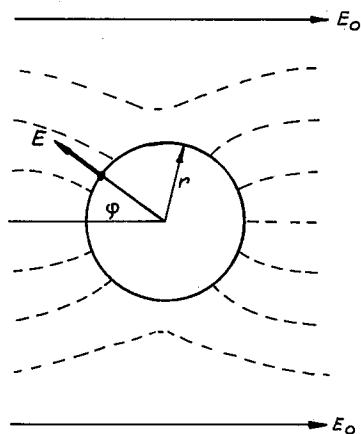
Zakres stosowalności równania Stokesa (5) i równania Oseena (7) przedstawiono na rys. 2 [7].

### 3. ELEKTRYZOWANIE CZĄSTEK

W niektórych metodach badania aerozoli, cząstki są wstępnie elektryzowane. Wyróżnia się trzy podstawowe mechanizmy elektryzowania cząstek:

- elektryzowanie zderzeniowe,
- elektryzowanie dyfuzyjne,
- elektryzowanie wychwytowe.

Przy wyprowadzeniu wyrażeń na ładunek uzyskiwany przez cząstkę założono kulisty kształt cząstki oraz równomierny rozkład pola w obszarze elektryzacji. Natężenie pola elektrycznego na powierzchni przewodzącej sfery umieszczonej w jednorodnym polu elektrycznym zewnętrznym  $E_0$  (rys. 3) opisuje zależność [10]:



3. Linie sił pola elektrycznego na powierzchni przewodzącej kuli

$$E = 3E_0 \cos \varphi. \quad (9)$$

Jeśli cząstka posiada własny ładunek  $Q$  natężenie pola zmienia się [11]:

$$E = 3E_0 \cos \varphi - \frac{Q}{4\pi\epsilon_0 r^2}. \quad (10)$$

Ujemne jony znajdujące się w powietrzu, będą przyciągane w tym obszarze sfery gdzie wypadkowe pole jest dodatnie, tzn. dla  $\cos \varphi \rightarrow 1$ . Jeśli natężenie pola przy powierzchni cząstki osiągnie wartość ujemną proces elektryzacji zakończy się. Maksymalny ładunek uzyskany przez cząstkę wynosi:

$$Q = 12\pi\epsilon_0 E_0 r^2. \quad (11)$$

Dla cząstek nieprzewodzących o stałej dielektrycznej  $\epsilon$ , zgromadzony ładunek jest mniejszy [11—14]:

$$Q = 4\pi\epsilon_0 E_0 r^2 \left( 1 + 2 \frac{\epsilon_r - 1}{\epsilon_r + 2} \right) \quad (12)$$

Elektryzowanie zderzeniowe jest dominujące dla cząstek większych od około 1  $\mu\text{m}$ . Dla mniejszych cząstek zaczyna przeważać elektryzowanie dyfuzyjne [11, 12, 13, 15].

Potencjał ujemnie naładowanej cząstki wynosi:

$$V = - \frac{ne^2}{4\pi\epsilon_0 r}, \quad (13)$$

gdzie  $n$  jest ilością elementarnych ładunków  $e$  zgromadzonych na cząstce.

Koncentracja jonów w pobliżu cząstki

$$N = N_0 \exp \left( - \frac{ne^2}{4\pi\epsilon_0 r k T} \right), \quad (14)$$

przy czym:  $N_0$  — koncentracja jonów w nieskończonej odległości od cząstki,  $k$  — stała Boltzmanna,  $T$  — temperatura. W elektryzowaniu dyfuzyjnym jony mogą osiągać powierzchnię cząstki ze wszystkich kierunków. Ilość zderzeń jonów z cząstką dana jest równaniem:

$$\frac{dn}{dt} = \pi r^2 \bar{v} N_0 \exp \left( - \frac{ne^2}{4\pi\epsilon_0 r k T} \right), \quad (15)$$

w którym  $\bar{v}$  — prędkość średniokwadratowa jonów.

Rozwiązanie tego równania prowadzi do wyrażenia na ładunek uzyskiwany przez cząstkę:

$$Q = \frac{4\pi\epsilon_0 r k T}{e} \ln \left( 1 + \frac{r \bar{v} N_0 e^2}{4\pi k T} t \right). \quad (16)$$

Zakłada się przy tym, że wszystkie jony, które osiągną powierzchnię cząstki pozostaną na niej, powiększając jej ładunek.

Z równania (16) wynika, że zgromadzony ładunek zmierza do nieskończoności przy nieskończonym czasie przebywania w obszarze zjonizowanym, tzn. z wyrażenia tego nie wynika ograniczenie procesu elektryzowania. Różni autorzy czynią dodatkowe założenia mające przezwyciężyć tę trudność, lecz mają one raczej charakter arbitralny

niż teoretycznie uzasadniony. Np. w [11] założono, że stan ustalony elektryzacji osiągany jest gdy czynnik

$$\frac{ne^2}{4\pi\epsilon_0rkT} \approx 5. \quad (17)$$

Osiągnięty ładunek wynosi wówczas

$$Q = \frac{20\pi\epsilon_0rkT}{e}. \quad (18)$$

Brak jednak teoretycznych podstaw do takich założeń.

Dla dużych prędkości cząstek aerozolu może zachodzić zjawisko wychwytywania jonów znajdujących się na torze ruchu cząstki. Zachodzi to wówczas, gdy prędkość cząstki przewyższa prędkość dryfu jonu w odpychającym polu cząstki:

$$v_i = \mu_i E \leq v_p \quad (19)$$

przy czym:  $v_i$ ,  $v_p$  — prędkość jonu i cząstki,  $\mu$  — ruchliwość jonu,  $E$  — natężenie pola elektrycznego przy powierzchni cząstki.

Przyjmując, że natężenie pola cząstki dane jest zależnością (9), dla  $\varphi = \pi/2$  maksymalny ładunek uzyskany przez cząstkę wyniesie:

$$Q = \frac{4\pi\epsilon_0r^2v_p}{\mu_i} \quad (20)$$

#### 4. ELEKTROSTATYCZNE ZAWIESZENIE CZĄSTEK

Jedną z najlepiej znanych metod w analizie cząstek jest metoda opracowana przez Millikana. W metodzie tej naelektryzowane cząstki zawieszone są w polu elektrostatycznym równoległym do sił ciężkości. Gdy ciężar cząstki zostaje zrównoważony przez siłę wywieraną przez pole elektrostatyczne (3) możliwe jest określenie ładunku właściwego cząstki

$$\frac{Q}{m} = \frac{g}{E} \quad (21)$$

Aby wyznaczyć masę i wielkość cząstki należy znać jej ładunek. Swobodne zawieszenie cząstek w tym układzie nie umożliwia niezależnego określenia wielkości i ładunku cząstki. Możliwe to jest dopiero po wykonaniu modyfikacji w układzie elektrod dokonanej przez Straubela [16—19]. Pomiedzy dwiema elektrodami kondensatora płaskiego zasilanego ze źródła napięcia stałego umieszczona jest trzecia elektroda zasilana napięciem zmiennym. W środku tej elektrody wykonany jest niewielki otwór (rys. 4). Urządzenie to umożliwia zawieszenie cząstek o masie dochodzącej do  $10^{-8}$  kg, podczas gdy w kondensatorze Millikana nie można utrzymać cząstek cięższych od  $10^{-15}$  kg.

Cząstki o stosunku  $Q/m$  nie przekraczającym pewnej wartości pozostają stabilne w środku otworu elektrody środkowej. Po przekroczeniu pewnej wartości ładunku właściwego cząstka zaczyna oscylować prostopadle do elektrody. Ruch cząstki opisany jest równaniem różniczkowym Mathieu [16]:

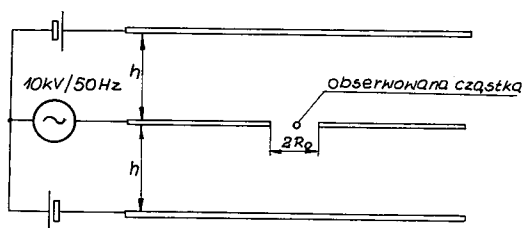
$$\ddot{x} + 2\gamma\dot{x} + (\alpha - 2\beta\cos\omega t)x + k = 0, \quad (22)$$

w którym:  $x$  — współrzędna położenia cząstki,

$$\alpha = \frac{-16QU_{\sim}}{\pi\omega^2 h R_0 m}, \quad (23)$$

$$\beta = \frac{8QU_{\sim}}{\pi\omega^2 h R_0 m}, \quad (24)$$

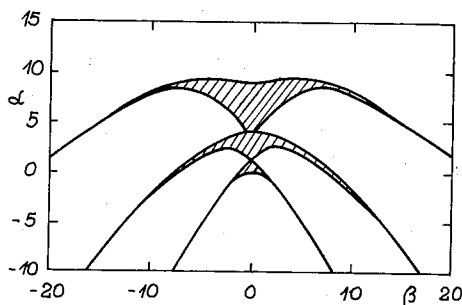
$$\gamma = \frac{6\pi\eta r}{m\omega} \quad (25)$$



4. Kondensator Straubela do badania wielkości cząstek

oraz  $Q, m, r$  — ładunek, masa i promień cząstki,  $\omega$  — częstotliwość napięcia zmiennego,  $R_0$  — promień otworu w elektrodzie środkowej,  $\eta$  — lepkość powietrza.

W równaniu (22) założono, że w otworze znajduje się tylko jedna cząstka. Dla wyznaczenia masy cząstki mają znaczenie tylko stabilne i okresowe rozwiązania równania (22), które zaznaczono na diagramie Strutta (rys. 5) [16, 17, 20] przy czym zarówno  $\alpha$  jak i  $\beta$  muszą być jednocześnie różne od zera.

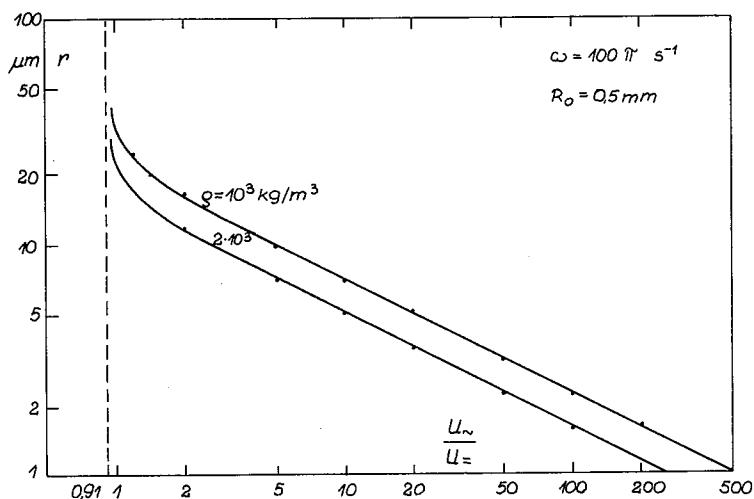


5. Diagram Strutta wg [16]

Wzrost wartości  $\beta$  przy podwyższaniu napięcia zmiennego aż do wartości

$$\beta = \sqrt{0,8245 + 4\gamma^2} \quad (26)$$

nie powoduje zmiany stabilnego położenia cząstki. Przekroczenie tej wartości powoduje oscylacje cząstki z amplitudą zależną od  $\beta$  i częstotliwością będącą połową częstotliwości napięcia zasilającego. Wartość  $\beta$ , dla której zaczynają się oscylacje umożliwia określenie ładunku właściwego  $Q/m$  z dokładnością lepszą od  $10^{-4}$ . Niezależne określenie wielkości cząstki jest możliwe tylko dla cząstek o promieniu mniejszym od około 25  $\mu\text{m}$ , gdy drugi składnik w (26) staje się porównywalny z wartością 0,8245. Wartość stosunku napięć  $U_{\sim}/U_{=}$ , przy której zaczynają się oscylacje można wyznaczyć z porównania (24) i (26).



6. Zależność promienia cząstki od stosunku napięć  $U_{\sim}/U_{=}$  w doświadczeniu Straubela

W wyrażeniu (24) stosunek  $Q/m$  może być wyeliminowany z równania Millikana (21) w którym  $U_{=}$  jest napięciem dającym stabilne położenie cząstki (w tym przypadku  $U_{\sim} = 0$ ). Podstawiając tak przekształcone wyrażenie na  $\beta$  do (26) otrzymuje się równanie

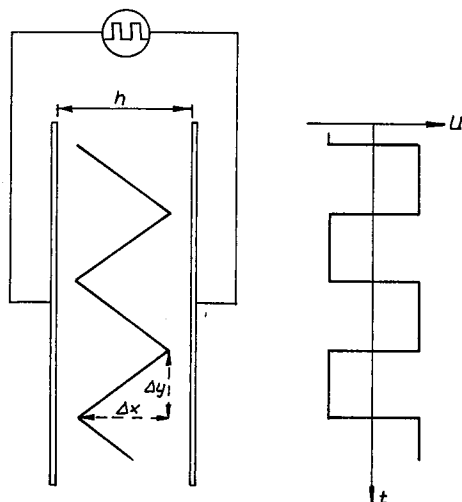
$$\left( \frac{16g}{\pi\omega^2 R_0} \right)^2 \left( \frac{U_{\sim}}{U_{=}} \right)^2 = 0,8245 + 4 \left( \frac{9\eta}{2\rho\omega r^2} \right)^2, \quad (27)$$

z którego można wyznaczyć promień cząstki. Zależność ta przedstawiona została na rys. 6. Pierwszy czynnik po lewej stronie (27) nie wpływa na zakres stosowności metody lecz tylko na nachylenie prostej na rys. 6, tak, że przyjęto iż jest on równy jedności (co odpowiada średnicy otworu około 1 mm). Obliczenia przeprowadzono dla częstotliwości 50 Hz. Nieznaczne rozszerzenie zakresu można uzyskać zmniejszając częstotliwość, lecz sposób ten jest ograniczony ze względu na stabilność rozwiązań.

## 5. METODY SEDYMENTACYJNE

W tej grupie metod badane są tory swobodnie opadających cząstek aerozolu w stałym lub zmiennym polu elektrycznym poprzecznym do pola grawitacyjnego. Z kształtu toru zakreślonego przez cząstkę (która jest fotografowana) wyznaczyć można zarówno ładunek jak i masę (wielkość) cząstki.

Metoda ta wprowadzona przez Wellsa i Gerke'a [21] była modyfikowana przez wielu autorów. W metodzie Wellsa i Gerke'a elektrody kondensatora płaskiego zasilane były napięciem prostokątnym (rys. 7). Tory cząstek miały kształt trójkątny. Z



7. Tory cząstek w polu o zmienności prostokątnej

równowagi siły grawitacji (1) i siły oporu powietrza (2), na podstawie składowej pionowej ruchu cząstki (5) można wyznaczyć jej wielkość:

$$r = \sqrt{\frac{9\eta f \Delta y}{2g}} \quad (28)$$

gdzie  $f$  jest częstotliwością napięcia zasilającego. Okres drgań cząstki w kierunku  $x$  i  $y$  jest równy  $1/2f$ . Metoda ta umożliwia pomiar wielkości cząstek większych od  $0,5 \mu\text{m}$ .

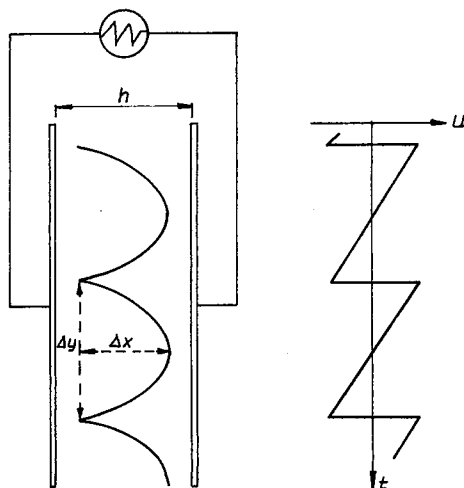
Mirgel [22] zastosował piłokształtne napięcie zasilające. Ten kształt pola powoduje, że cząstki zakreślają tory paraboliczne (rys. 8). Z równowagi siły grawitacji (1) i siły Stokesa (2), wyznaczyć można wielkość cząstek:

$$r = \sqrt{\frac{9\eta f \Delta y}{2g}}, \quad (29)$$

przy czym:  $\Delta y$  — składowa pionowa drogi cząstki za okres napięcia zasilającego.

Stałe pole elektrostatyczne zostało zastosowane przez Kunkela i Hansena [23, 24]. Efekt ruchu na kliszy fotograficznej uzyskiwany był dzięki zastosowaniu pulsującego

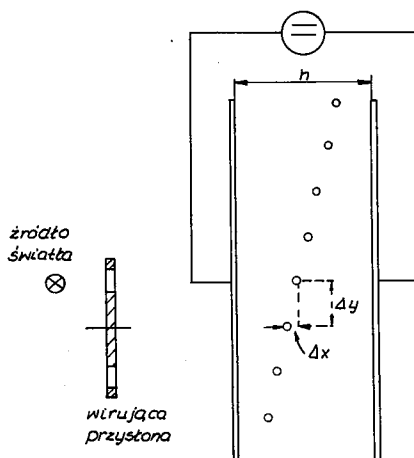
światła modulowanego za pomocą wirującej tarczy. Na filmie uzyskano ciąg punktów (rys. 9). Sposób ten umożliwia pomiar cząstek o promieniu większym od  $1\ \mu\text{m}$ . Dla mniejszych cząstek należy uwzględnić współczynnik korekcyjny Cunninghama. Dokładność metody nie przekracza 1%.



8. Tory cząstek w polu o zmienności piłokształtnej

Z równania Stokesa i średniego odstepu pomiędzy punktami na kliszy można wyznaczyć promień obserwowanej cząstki:

$$r = \sqrt{\frac{9\eta f \Delta y}{2qgMv}}, \quad (30)$$



9. Ślady cząstek opadających w stałym polu elektrycznym oświetlonych impulsowo

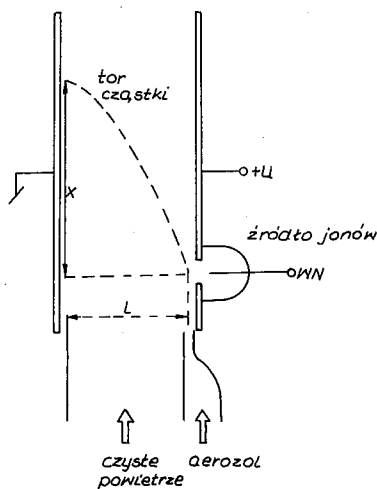
przy czym:  $f$  — częstotliwość impulsów światła,  $v$  — liczba przedziałów pomiędzy punktami na kliszy wziętymi do obliczeń na drodze  $\Delta y$ ,  $M$  — powiększenie kamery.

We wszystkich tych metodach przytoczone proste zależności na wielkość cząstki są słuszne tylko dla cząstek kulistych. Wielkość cząstki o innym kształcie może być wyznaczona tylko w przybliżeniu a błąd pomiaru może przekraczać 10%.

Zaletą tych metod oraz metody Straubela jest możliwość niezależnego określenia masy (wielkości) oraz ładunku niesionego przez cząstkę, co ma istotne znaczenie w przypadku badania aerozolu naelektryzowanego. Żadna z pozostałych metod nie daje takiej możliwości.

## 6. KLASYFIKACJA ELEKTROSTATYCZNA

W tej grupie metod, w odróżnieniu od opisanych rozdz. 5, cząstki nie opadają swobodnie lecz poruszają się w strumieniu powietrza. Tory cząstek nie są fotografowane, lecz bada się całkowity prąd płynący do elektrody zbiorczej, pochodzący od cząstek rozseparowanych w polu elektrostatycznym. W miejsce elektrody zbiorczej (kolektora) umieszcza się często płytki mikroskopowe.



10. Urządzenie Rohmanna do pomiaru wielkości cząstek

Pierwszy ideę klasyfikacji (separacji) elektrostatycznej zastosował Rohmann [25]. Przyrząd składał się z dwóch równoległych płyt tworzących kondensator płaski, pomiędzy którymi przepływał gaz z zawieszonymi cząstkami aerozolu (rys. 10). Pomiędzy płytami utrzymywana jest wysoka różnica potencjałów. Przy brzegu jednej z elektrod znajduje się dysza, przez którą wpływa badany aerozol. Tuż za dyszą

umieszczony jest układ elektryzacji aerozolu. Ładunek uzyskany przez każdą cząstkę w obszarze elektryzacji wynosi:

$$Q = \frac{4\pi \left(1 + 2 \frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right) \varepsilon_0 r^2 U}{h \left(1 + \frac{4\varepsilon_0 U b v_a}{neh}\right)}, \quad (31)$$

przy czym:  $\varepsilon_r$  — względna przenikalność elektryczna cząstki,  $n$  — ilość jonów przepływających przez obszar elektryzacji w jednostce czasu,  $e$  — ładunek elementarny,  $b$  — długość szczeliny elektryzatora (poprzecznie do kierunku przepływu aerozolu),  $v_a$  — średnia prędkość powietrza,  $h$  — odległość między okładkami.

Składowa prędkości cząstek prostopadła do elektrod wynosi

$$v_y = \frac{QU}{6\pi\eta rh}. \quad (32)$$

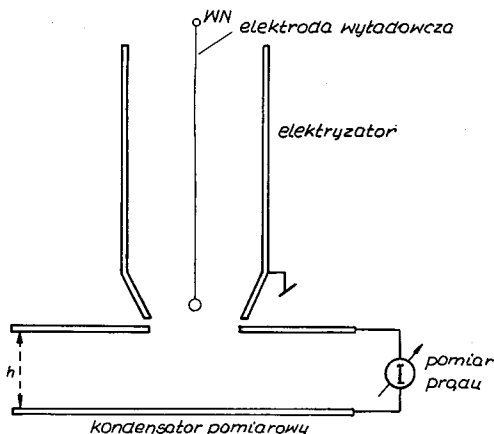
Odległość  $x$  jaką dana cząstka przebędzie w strumieniu powietrza od wejścia do punktu, w którym osiadzie na elektrodzie

$$x = \frac{3\eta v_a l h^2 \left(1 + \frac{4\varepsilon_0 b v_a U}{neh}\right)}{2\varepsilon_0 r U^2 \left(1 + 2 \frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right)}, \quad (33)$$

gdzie:  $l$  — odległość wlotu aerozolu do elektrody zbiorczej.

Cząstki mogą być osadzone na szkiełkach mikroskopowych lub można mierzyć prąd z małego wycinka elektrody zbiorczej w zależności od napięcia przyłożonego do elektrod, uzyskując względny rozkład wielkości cząstek.

Sposób opisany przez Drozina i La Mer'a [26] opiera się na pomiarze prądu niesionego przez naładowane cząstki opadające równoległe do linii sił pola elek-



11. Urządzenie Drozina — La Mera do pomiaru gęstości rozkładu cząstek

trycznego w polu kondensatora płaskiego (rys. 11). Aerosol elektryzowany jest w kondensatorze cylindrycznym, w którym wytworzono wyładowanie ulotowe, przez czas około 1 minuty. Następnie aerosol wpuszczany jest do kondensatora płaskiego. Pole w kondensatorze płaskim jest na tyle silne aby można było pominąć siły grawitacji. W tym przypadku czas osiadania cząstek o promieniu  $r$  wynosi:

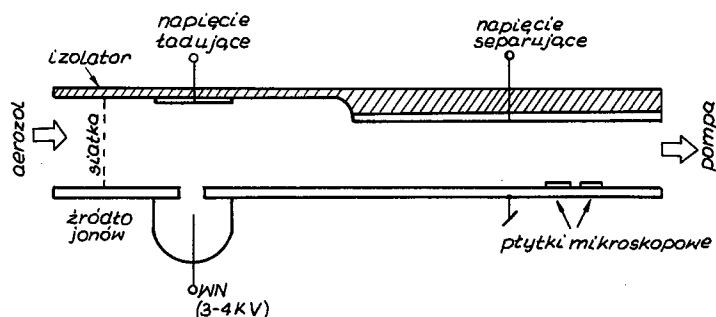
$$t = \frac{6\pi\eta hr}{QE}. \quad (34)$$

Natężenie prądu płynącego do dolnej elektrody jest więc funkcją czasu. Jeśli znana jest zależność ładunku uzyskiwanego przez cząstkę od jej promienia, z czasowej zależności natężenia prądu można wyznaczyć gęstość rozkładu promieni cząstek aerosolu. W przypadku elektryzacji opisanej zależnością (12) promień cząstek, które dotrą do kolektora po czasie  $t$  wynosi:

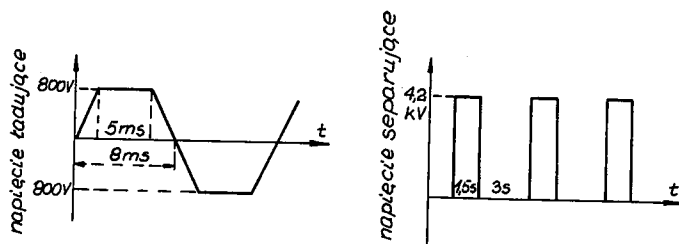
$$r = \frac{\eta h^2(2 + \epsilon_r)}{2\epsilon_0 \epsilon_r U t}. \quad (35)$$

Należy zaznaczyć, że w przypadku, gdyby ładunek uzyskiwany przez cząstkę był proporcjonalny do jej promienia, czas osiadania cząstek byłby niezależny od ich wielkości.

Do badania cząstek submikronowych służy przyrząd opracowany przez Liu, Whitby i Yu [15]. Urządzenie utworzone jest przez płaski kanał podzielony na dwie części: obszar elektryzacji i obszar separacji (rys. 12). W obszarze elektryzacji cząstki



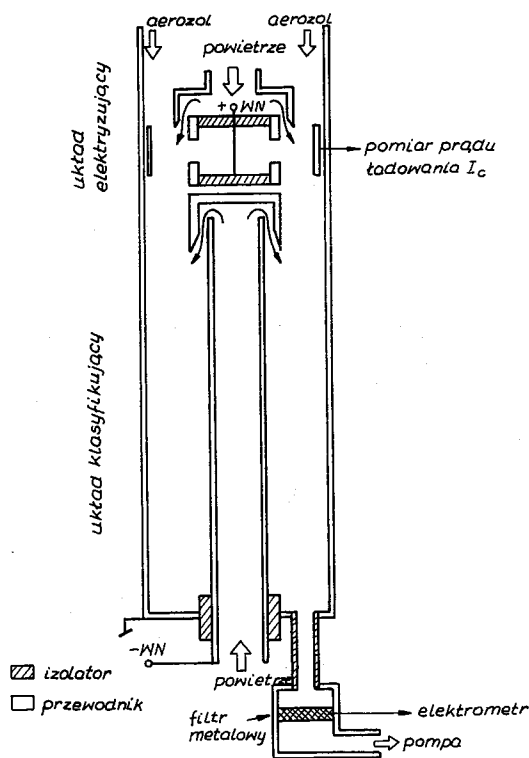
12. Urządzenie do pomiaru rozkładu cząstek submikronowych



13. Przebiegi napięcia ładującego i klasyfikującego w układzie z rys. 12.

ładowane są przez jony dodatnie wytworzone w wyładowaniu ulotowym pod jedną z elektrod. Aby uniknąć zakrzywienia toru cząstek elektrody elektryzatora zasilane są napięciem przemiennym o kształcie jak na rys. 13. Kondensator separatora zasilany jest również napięciem impulsowym. W czasie gdy napięcie wynosi zero cała objętość kondensatora płaskiego wypełniona jest naelektryzowanym aerozolem. W czasie impulsu napięcia cząstki aerozolu osiadają równomiernie na dolnej, uziemionej elektrodzie, na której umieszcza się szkiełka mikroskopowe. Metoda ta jest efektywna dla cząstek o wielkości od około  $0,01 \mu\text{m}$  do około kilku  $\mu\text{m}$ .

Analizator aerozolu opisany przez Whitby'ego i Clarka [27] oraz przez Liu i Pui [28—30] składa się z dwóch koncentrycznych cylindrów, pomiędzy którymi przepływa badany aerozol. Cząstki aerozolu są elektryzowane w górnej części urządzenia przez prąd wyładowania koronowego (rys. 14). Te cząstki, które nie osiadają na



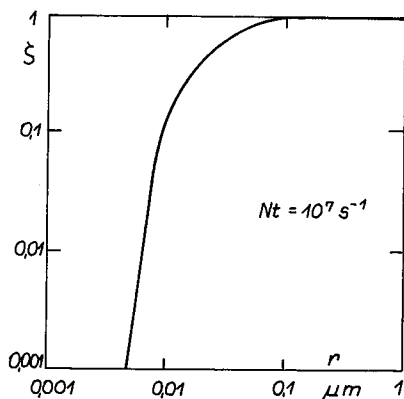
14. Układ do pomiaru dystrybucji rozkładu cząstek submikronowych

wewnętrznym cylindrze kondensatora są odpompowywane przez otwór w dolnej części kolumny. Cząstki te wychwytywane są przez filtr metalowy. Prąd niesiony przez te cząstki mierzony jest za pomocą elektrometru. Zmieniając napięcie między elektrodami i mierząc prąd płynący przez filtr można wyznaczyć ruchliwość cząstek oraz dystrybucję wielkości cząstek. Urządzenie wymaga kalibrowania za pomocą aerozolu monodispersyjnego. Dla aerozolu monodispersyjnego o koncentracji czą-

stek  $N_p$ , z których każda posiada  $n$  ładunków elementarnych, prąd mierzony przez elektrometr wynosi:

$$I = \dot{V}_a n e N_p \xi \quad (36)$$

przy czym:  $\dot{V}_a$  — objętościowe natężenie przepływu aerozolu,  $\xi$  — empiryczny współczynnik korekcyjny uwzględniający straty aerozolu spowodowane niecałko-

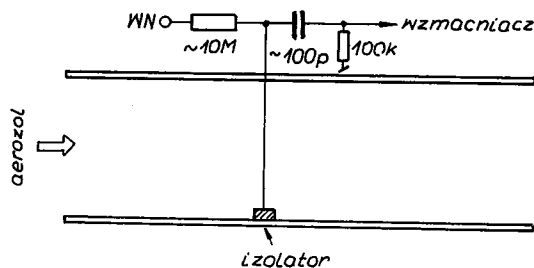


15. Współczynnik korekcyjny dla układu z rys. 14 (wg [28]).

witym naelektryzowaniem cząstek oraz dyfuzją cząstek. Czynniki  $\xi$  jest praktycznie równy jedności dla cząstek większych od  $0,5 \mu\text{m}$  lecz może zmniejszyć się od wartości 0,001 dla cząstek o promieniu  $5 \text{ nm}$  (rys. 15). Rozkład wielkości cząstek może być wyznaczony z charakterystyki prądowo-napięciowej na podstawie zależności (36).

## 7. SONDY DRUTOWE

Cząstki przewodzące mogą być badane za pomocą sond drutowych znajdujących się na wysokim potencjale elektrycznym. Cząstka uderzając w cienki drut odbiera z niego ładunek wytwarzając impuls w obwodzie elektrycznym (rys. 16) [31—33]. Drut używany w tego rodzaju sondach powinien mieć średnicę mniejszą niż  $0,1 \text{ mm}$ .



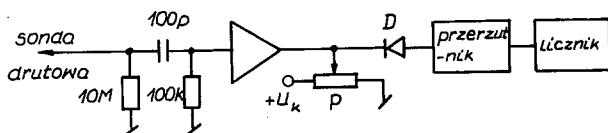
16. Sonda drutowa z oddawaniem ładunku

Zależność empiryczna pomiędzy amplitudą impulsu a promieniem cząstki podana została w [32]:

$$U_p = AUr^{1,6}, \quad (37)$$

gdzie:  $U$  — potencjał drutu,  $A$  — stała zależna od średnicy drutu. Krople wody uderzając w drut wytwarzają impulsy o czasie narastania około  $1 \mu s$ . Napięcie zasilające powinno być możliwie duże aby uzyskać jednoznaczne naładowanie kropeł. Jednak praktycznie jest ono ograniczone do około 4 kV ze względu na występujące wyładowanie koronowe.

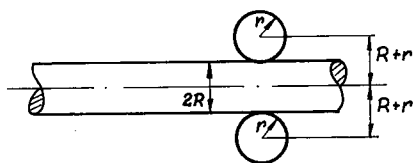
Uzyskane impulsy mogą być klasyfikowane przez dyskryminator okienkowy. Uzyskuje się wówczas gęstość rozkładu wielkości kropeł. Zamiast układów klasyfikacji Gardiner zaproponował prosty układ do wyznaczania dystrybucji rozkładu [33] (rys. 17). Dioda  $D$  jest spolaryzowana zaporowo potencjałem pochodzącym z



17. Układ do pomiaru dystrybucji dla sondy drutowej

potencjometru  $P$ , który reguluje napięcie od zera do maksymalnej wartości napięcia wyjściowego wzmacniacza. Impulsy, których amplituda przekracza napięcie polaryzacji, uruchamiają przerzutnik przesyłający impulsy do licznika.

Dla kropeł o średnicy dużo większej od średnicy drutu należy wprowadzić czynnik korekcyjny uwzględniający „powiększenie” czynnego przekroju drutu (rys. 18). Dla



18. Sposób określania przekroju czynnego sondy drutowej

sondy o średnicy  $2R$  i długości  $L$  powierzchnia czynna dla kropli o średnicy  $2r$  wynosi  $2L(R+r)$ . Aby uniezależnić postać uzyskanego rozkładu od wielkości kropli należy wynik pomiaru dla danej wielkości kropeł pomnożyć przez czynnik

$$\xi = \frac{R}{R+r}. \quad (38)$$

Metoda ta umożliwia pomiar cząstek o promieniu większym od  $1 \mu m$ . Częstość zliczeń może dochodzić do 200 000/min. Przyrząd wymaga kalibracji za pomocą cząstek o znanej wielkości.

Guyton [31, 34] wykorzystał sondę drutową do pomiaru ładunku niesionego przez naelektryzowane cząstki. Cząstki w zetknięciu z uziemionym drutem oddają ładunek wytwarzając impuls w obwodzie elektrycznym. Cząstki elektryzowane są w czasie przejścia z dużą prędkością przez cienką szklaną dyszę. Ładunek uzyskiwany przez cząstkę jest proporcjonalny do jej powierzchni. Napięcie impulsu jest kwadratową funkcją promienia cząstki

$$U = A \frac{r^2}{C}, \quad (39)$$

przy czym:  $C$  — pojemność wejściowa układu elektrycznego,  $A$  — stała zależna od sposobu elektryzacji cząstek. Metoda ta jest efektywna dla cząstek o średnicy większej od  $3 \mu\text{m}$ .

Zastosowanie obu metod do pomiaru rozkładu wielkości kropeł cieczy jest ograniczone. W pierwszym przypadku z powodu zwilżania i osiadania kropli cieczy na drucie. Impulsy pochodzić będą wówczas od dużych kropli oderwanych z drutu przez strumień powietrza. W drugim przypadku, przejście aerozolu przez dyszę zmieni rozkład wielkości kropeł.

## 8. SONDY IGŁOWE

Sondy tego rodzaju składają się z dwóch metalowych igieł ustawionych naprzeciw siebie końcami i tworzących niewielką szczelinę. Jeśli kropla przewodzącej cieczy dotknie jednocześnie końców obu igieł, zamknie obwód elektryczny i wytworzy impuls elektryczny [35, 36]. Zmieniając odległość między końcami igieł można uzyskać dystrybuantę rozkładu wielkości kropeł.

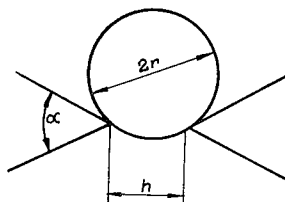
Liczba impulsów uzyskanych w jednostce czasu wyraża się zależnością:

$$H(h) = \int_h^{D_{\max}} K(h,r) f(r) dr, \quad (40)$$

w której:  $h$  — odległość między końcami igieł,  $r$  — promień kropli,  $f(r)$  — funkcja częstości rozkładu wielkości kropli przechodzących przez badaną objętość. Funkcja  $K(h,r)$  wyraża prawdopodobieństwo dotknięcia przez kroplę obu końców igieł.

Dla elektrod stożkowych (rys. 19):

$$K(h,r) = \frac{2r^2}{p} \left[ \beta + \text{tg}^{-1} \beta + \frac{\left(\frac{h}{2r}\right)^2 \cos \beta + \frac{h}{r}}{\sin \beta} \right], \quad (41)$$



19. Kropla zawierająca elektrody sondy igłowej

gdzie:  $P$  — pole powierzchni przekroju poprzecznego kanału pomiarowego,  $\beta = (\pi - \alpha)/2$ ,  $\alpha$  — kąt przy wierzchołku stożka. Równanie (41) słuszne jest dla  $2r > h$  i  $0 < \beta \leq \pi/2$ .

Jeśli wyznaczona zostanie funkcja  $H(h)$  gęstość rozkładu wielkości kropeł może być obliczona z (40) przy założeniu maksymalnej wielkości kropli  $r_{\max}$  [36].

## 9. SONDY PIEZOELEKTRYCZNE

Sondy piezoelektryczne wykorzystują:

1) generację impulsu elektrycznego przy uderzeniu cząstki w kryształ piezoelektryka [37],

2) zmianę częstotliwości własnej rezonansu kryształu przy zmianie jego masy przez osiadającą cząstkę [38—42].

W pierwszym przypadku cząstka uderzając w kryształ przekazuje mu swój pęd. Wygenerowane napięcie jest proporcjonalne do czwartej potęgi promienia cząstki. Jeśli znana jest prędkość i gęstość cząstek można obliczyć wielkość cząstek aerozolu. Kryształ piezoelektryka może być umieszczony w wolnym przepływie lub stanowić płytkę w impaktorze kaskadowym, podobnym do przedstawionego przez Friedlanda [43].

W drugiej metodzie piezoelektryk wykonany jest w postaci płytki o średnicy paru mm i grubości 0,1—0,2 mm pokrytej materiałem umożliwiającym przyklepienie się cząstek. Grubość warstwy jest rzędu  $1\ \mu\text{m}$ . Dla małej amplitudy oscylacji i małych zmian masy, zmiana częstotliwości rezonansowej jest liniową funkcją przyrostu masy  $\Delta m$ :

$$\Delta f = -Cf^2\Delta m \quad (42)$$

przy czym stała  $C$  zależy od właściwości kryształu.

Zmiany częstotliwości mierzone są metodą porównawczą. Sposób ten umożliwia pomiar cząstek o masach większych od  $10^{-14}$  kg, co dla gęstości  $10^3\ \text{kg/m}^3$  odpowiada cząstkom o promieniu rzędu  $1\ \mu\text{m}$ .

## ZAKOŃCZENIE

W tabeli poniższej zestawiono metody opisane w pracy i zakres ich zastosowań. W trzeciej kolumnie podano, czy metoda umożliwia pomiar wielkości pojedynczej cząstki, dystrybucyjny rozkład czy gęstości rozkładu wielkości cząstek. Wszystkie metody służące do pomiaru wielkości pojedynczych cząstek mogą być wykorzystane do pomiaru gęstości rozkładu po wykonaniu odpowiedniej ilości pomiarów. W tabeli pominięto tak ważny parametr opisujący przydatność metody jak graniczna koncentracja cząstek, gdyż na ogół nie jest on podawany przez autorów publikacji. Ostatnia kolumna odsyła do rozdziału, w którym dana metoda została opisana.

| Metoda                 | Zakres              | Zastosowanie                  | Rozdział |
|------------------------|---------------------|-------------------------------|----------|
| Straubela              | $< 25 \mu\text{m}$  | pojedyncze cząstki            | 4        |
| Wellsa i Gerke'a       | $> 0,5 \mu\text{m}$ | pojedyncze cząstki            | 5        |
| Mirgela                | $> 1 \mu\text{m}$   | pojedyncze cząstki            | 5        |
| Kunkela i Hansena      | $> 1 \mu\text{m}$   | pojedyncze cząstki            | 5        |
| Rohmanna               | $> 0,1 \mu\text{m}$ | gęstość rozkładu              | 6        |
| Drozina i La Mera      | $> 0,1 \mu\text{m}$ | gęstość rozkładu              | 6        |
| Liu i Pui              | $> 5 \mu\text{m}$   | dystribuanta rozkładu         | 6        |
| Liu, Whitby i Yu       | $> 10 \mu\text{m}$  | gęstość rozkładu              | 6        |
| Gardinera/Geista       | $> 1 \mu\text{m}$   | gęstość/dystribuanta rozkładu | 7        |
| Guytona                | $> 1 \mu\text{m}$   | gęstość/dystribuanta rozkładu | 7        |
| Sonda igłowa           | $> 5 \mu\text{m}$   | dystribuanta rozkładu         | 8        |
| Sonda piezoelektryczna | $> 1 \mu\text{m}$   | gęstość/dystribuanta rozkładu | 9        |

## BIBLIOGRAFIA

1. TSI — katalog
2. MALVERN — katalog
3. Z. Orzechowski: *Rozpylanie cieczy*. Warszawa: WNT, 1976
4. K. Leschonski, W. Alex, B. Koglin: *Teilchengrößenanalyse*. Chemie, Ingenieur, Technik, (1974), v. 46, No 1, 3, 7, 9, 11, 13, 15, 17, 19, 21, 23; (1975), v. 47, No 1, 3
5. W. Alex: *Prinzipien und Systematik der Zählverfahren in der Teilchengrößenanalyse*. Aufbereitungs-Technik No 2/1972 s. 105—11; No 3/1972 s. 168—82; No 10/1972 s. 639—52; No 11/1972 s. 723—32
6. A. Bürkholz: *Messmethoden zur Tropfengrößenbestimmung*. Chemie, Ingenieur, Technik, (1973), v. 45, No 1, s. 1—7
7. T. Gast: *Staubmesstechnik — eine Übersicht*. Technisches Messen ATM, (1978), No 12, s. 427—36
8. K. Leschonski: *Kennzeichnung disperser Systeme, Teilchengrößenanalyse*. Chemie, Ingenieur, Technik, (1973), v. 45, No 1, s. 8—18
9. B.J. Azzopardi: *Measurement of Drop Size*. Int. J. Heat and Mass Transfer, (1979), v. 22, No 9, s. 1245—75
10. E. Philippow: *Grundlagen der Elektrotechnik*. Leipzig, 1976
11. J. Böhm: *Electrostatic Precipitators*. Elsevier/SNTL, Praha, 1982
12. A.D. Moore (ed): *Electrostatics and its Application*. J. Wiley a. Sons, 1973
13. H.J. White: *Industrial Electrostatic Precipitation*. Pergamon Press, 1963
14. I.I. Inculat: *Particle Charging in DC Corona Field*. IEEE Trans. Electrical Insulation, (1982), v. EI-17, No 2, s. 168—71
15. B.Y.H. Liu, K.T. Whitby, H.H.S. Yu: *Electrostatic Aerosol Sampler for Light and Electron Microscopy*. Rev. Sci. Instr., (1967), v. 38, No 1, s. 100—2
16. H. Straubel: *Die Stabilisierung elektrisch geladener Teilchen in Wechselfeldern*. Acta Physica Austriaca, (1960), v. 13, No 2/3, s. 265—73

17. H. Straubel: *Bestimmung der Ladungsverteilung von Oberflächen mittels stabilisierter elektrisch geladener Teilchen*. Staub, Reinhaltung der Luft, (1968), v. 28, No 12, s. 506—8
18. H. Straubel: *A Non-Optical Method for Particle Sizing*. Atmospheric Pollution. 12-th International Colloquium, Paris, May 5—7, 1976, s. 471—6
19. H. Straubel: *Elektro-optische Messung von Aerosolen*. Technisches Messen ATM, (1981), v. 48, No 6, s. 199—210
20. A. Müller: *Theoretische Untersuchungen über das Verhalten geladener Teilchen in Sattelpunkten elektrischer Wechselfelder*. Annalen der Physik, (1960), v. 6, s. 206—20
21. P. V. Wells, R. H. Gerke: *An Oscillation Method for Measuring the Size of Ultramicroscopic Particles*. J. American Chemical Society, (1919), v. 41, s. 312
22. K. H. Mirgel: *Elektrostatistische Aufladung von Stäuben*. VDI-Berichte, (1957), v. 19, s. 49—57
23. W. B. Kunkel, J. W. Hansen: *A Dust Electricity Analyzer*. Rev. Sci. Instr., (1950), v. 21, No 4, s. 308—14
24. L. B. Loeb: *Static Electrification*. Springer Verlag, 1958
25. H. Rohmann: *Methode zur Messung der Grösse von Schwebeteilchen*. Z. Physik, (1923), v. 17, No 4/5, s. 253—65
26. V. G. Drozin, V. K. La Mer: *The Determination of the Particle Size Distribution of Aerosols by Precipitation of Charged Particles*. J. Colloid Sci., (1959), v. 14, s. 74—90
27. K. T. Whitby, W. E. Clark: *Electric Aerosol Particle Counting and Size Distribution Measuring System from the 0,015 to 1  $\mu\text{m}$  Size Range*. Tellus, (1966), v. 18, No 2, s. 573—86
28. B. Y. H. Liu, D. Y. H. Pui: *On the Performance of the Electrical Aerosol Analysis*. J. Aerosol Sci., (1975), v. 6, s. 249—64
29. B. Y. H. Liu, D. Y. H. Pui, A. Kapadia: *Electrical Aerosol Analyzer: History, Principle and Data Reduction*. Aerosol Measurement (Workshop), University of Florida 1979, s. 341—83
30. D. Y. H. Pui, B. Y. H. Liu: *Electrical Aerosol Analyzer: Calibration and Performance*. Aerosol Measurement (Workshop), University of Florida 1979, s. 384—99
31. A. C. Guyton: *Electronic Counting and Size Determination of Particles in Aerosols*. J. Industrial Hygiene and Toxicology, (1946), v. 28, No 4, s. 133—41
32. J. M. Geist, J. L. York, J. J. Brown: *Electronic Spray Analyzer for Electrically Conducting Particles*. Industrial and Engineering Chemistry, (1951), v. 43, No 6, s. 1371—7
33. J. A. Gardiner: *Measurements of the Drop Size Distribution in Water Sprays by an Electrical Method*. Instrument Practice, April 1964, s. 353—6
34. F. T. Gucker Jr, C. T. O'Konski: *Electronic Methods of Counting Aerosol Particles*. Chemical Rev., (1949), v. 44, s. 373—88
35. J. J. Picot, T. D. Tomey, J. B. Smedley, D. D. Kristmanson: *Spray Atomizer Droplet Spectrum Measurement*. Canadian J. Chemical Engineering, (1980), v. 58, No 6, s. 314—7
36. B. Durkiewicz, A. Adamkiewicz, J. Kowalski: *Pomiar średnic kropel wody oraz określenie ich rozkładu w przepływie powietrza mgłowego*. Zeszyty Naukowe Wyższej Szkoły Marynarki Wojennej, (1980), v. 21, No 4/67, s. 75—88
37. R. Davies: *A Review of the Methods for the Particle Size Analysis of Droplets, Sprays and Mists*. Partikelmesstechnik (Symposion), Nürnberg 1975, s. 115—34
38. T. E. Carpenter, D. L. Brechley: *A Piezoelectric Cascade Impactor for Aerosol Measuring*. American Industrial Hygiene Association J. Aug. 1972, s. 503—10
39. R. L. Chan: *An Instrument for the Direct Measurement of Particulate Mass*. J. Aerosol Sci., (1970), v. 1, s. 111—4
40. J. G. Olin, G. J. Sem, D. L. Christenson: *Piezoelectric Electrostatic Aerosol Mass Concentration Monitor*. American Industrial Hygiene Association J. April 1971, s. 209—20
41. D. C. Woods: *Measurement of Particulate Aerosol Mass Concentration Using a Piezoelectric Crystal Microbalance*. Aerosol Measurement (Workshop), University of Florida 1979, s. 119—30
42. G. J. Sem, P. S. Daley: *Performance Evaluation of a New Piezoelectric Aerosol Sensor*. Aerosol Measurement (Workshop), University of Florida 1979, s. 672—86
43. S. K. Friedlander: *Smoke, Dust and Haze*. Fundamentals of Aerosol Behavior. J. Wiley, 1977

A. JAWOREK

## ELECTRICAL METHODS OF AEROSOL MEASUREMENT

## Summary

The electrical methods for size measurement of a single particle and size distribution of aerosol particles are reviewed. The Straubel condenser, electrostatic sedimentation methods, electrostatic precipitators, wire and needle probes and piezoelectric probes are described. The principle of operation and range of application of each of the methods are given.

# Zależności między parametrami systemu automatycznego prowadzenia pociągu APP i systemu przekazywania informacji z toru do pociągu

HENRYK KARBOWIAK

*Wydział Elektryczny — Instytut Elektroenergetyki Trakcja Elektryczna, Politechnika Łódzka*

*Otrzymano 1989.04.10*

*Autoryzowano do druku 1989.07. 01*

Obecnie prowadzone są intensywne prace nad systemami automatycznego prowadzenia pociągu. W artykule przedstawiono próby wprowadzenia określonych parametrów charakteryzujących oba systemy i wyznaczenia korelacji między tymi parametrami. W opracowaniu wprowadzono parametr; droga, którą przejeżdża pociąg przy odbiorze 1 telegramu —  $S_T$  charakteryzujący system automatycznego prowadzenia pociągu (APP). Następnie wyprowadzono zależności między parametrem systemu APP, prędkością pociągu oraz parametrami systemu przekazywania informacji, szybkością modulacji i liczbą bitów w telegramie.

## 1. PARAMETRY SYSTEMU AUTOMATYCZNEGO PROWADZENIA POCIĄGU — APP

Systemy automatycznego prowadzenia pociągu — APP [1] pracują wykorzystując informacje przekazywane z toru do pojazdu trakcyjnego za pomocą obwodu transmisyjnego. Zakres automatyzacji prowadzenia pociągu decyduje o liczbie przekazywanych informacji w formie telegramu.

Liczba telegramów odbieranych przez system APP zależy w ogólności od prędkości pociągu i szybkości transmisji informacji. Aby wyznaczyć zależność między tymi parametrami należy wprowadzić wielkość określającą niezbędną liczbę telegramów jaką powinien otrzymać system APP w odniesieniu do jednostki drogi lub czasu. Przez telegram rozumiany jest zestaw informacji niezbędnych do prowadzenia pociągu wraz z elementami zabezpieczającymi i określającymi początek i koniec przekazywanej informacji.

Jeżeli przyjąć za jednostkę odniesienia drogę, to wielkość charakteryzująca system APP jest określana jako liczba telegramów przypadająca na jednostkę przejechanej drogi. Wygodniej jest jednak korzystać z jej odwrotności, tj. przyjąć drogę jaką musi przejechać pociąg, aby odebrać 1 telegram. Długość drogi, na której ma być odebrany 1 telegram zależna jest od prędkości pociągu i czasu trwania telegramu.

Jeżeli zaś przyjąć za jednostkę odniesienia czas, to wielkością charakteryzującą system APP jest liczba telegramów przypadająca na jednostkę czasu. W tym

przypadku liczba telegramów odebranych w jednostce czasu zależy od czasu trwania telegramu, a nie zależy od prędkości pociągu. Obie rozważane wyżej wielkości charakterystyczne dla systemu APP można zatem uznać za równoważne. Biorąc jednak pod uwagę, że wielkość odniesienia do jednostki drogi uwzględnia również prędkość pociągu, za celowe należy uznać przyjęcie za podstawowy parametr systemu automatycznego prowadzenia pociągu — APP wielkość określoną jako drogę — oznaczoną dalej  $S_T$  — przy przejechaniu której zostanie odebrany jeden telegram.

Urządzenia pojazdowe APP muszą otrzymywać telegramsy tak często, aby w czasie odbioru telegramu droga przebyta przez pociąg przy jeździe z największą prędkością  $V_{max}$  nie powodowała wydłużenia drogi hamowania w stopniu mogącym wpływać na zmniejszenie bezpieczeństwa ruchu.

Parametrem charakteryzującym system automatycznego prowadzenia pociągu — APP jest, z interesującego punktu widzenia, droga jaką przejeżdża pociąg w czasie odbioru jednego telegramu —  $S_T$ . Po uwzględnieniu prędkości maksymalnej pociągu —  $V_{max}$  otrzymujemy wymagany czas trwania telegramu  $\tau_T$ .

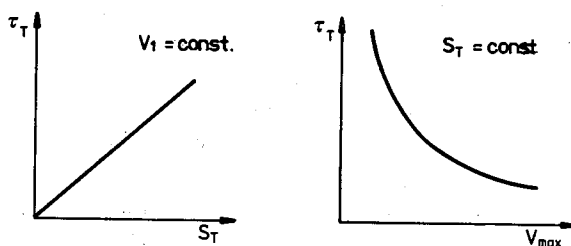
Przyjmując długość drogi przypadającej na jeden telegram  $S_T$  oraz znając prędkość maksymalną pociągu —  $V_{max}$  można wyznaczyć liczbę telegramów przypadających na jednostkę czasu  $W_T$  np. sekundę.

$$W_T = \frac{V_{max}}{S_T} \quad (1)$$

wygodniej jest operować pojęciem czas trwania telegramu  $\tau_T$ , który jest odwrotnością liczby telegramów przypadających na jednostkę czasu.

$$\tau_T = \frac{1}{W_T} = \frac{S_T}{V_{max}} [\text{s}]. \quad (2)$$

Zależności czasu trwania telegramu  $\tau_T$  od drogi przejeżdżanej przez pociąg w czasie odbioru 1 telegramu  $S_T$  i prędkości maksymalnej pociągu  $V_{max}$  pokazano na rys. 1.



Rys. 1. Zależność czasu trwania telegramu  $\tau_T$  od drogi przejeżdżanej przez pociąg w czasie odbioru jednego telegramu (rys. 1a) i prędkości pociągu  $V$  (rys. 1b)

Z rys. 1 wynika, że ze wzrostem przyjętej drogi przejeżdżanej przez pociąg przy odbiorze 1 telegramu czas trwania telegramu może być większy. Przy wzroście prędkości pociągu czas trwania telegramu ulega zmniejszeniu, a więc przy przekazywaniu tej samej liczby informacji i przy tym samym zabezpieczeniu telegramu musi następować zwiększenie szybkości modulacji.

## 2. PARAMETRY SYSTEMU PRZEKAZYWANIA INFORMACJI Z TORU DO POJAZDU

Systemy przekazywania informacji z toru do pojazdu z punktu widzenia współpracy z systemem APP charakteryzuje parametr określający szybkość modulacji  $v$ . Analizując zagadnienie szczegółowiej można wykazać, że szybkość modulacji zależy od częstotliwości  $F_0$  sygnału nośnego i liczby  $n$  okresów sinusoidy przypadających na jeden bit.

$$v = \frac{F_0}{n}. \quad (3)$$

W dalszych rozważaniach system przekazywania informacji charakteryzowany będzie parametrem szybkości modulacji  $v$ .

Zależność czasu trwania telegramu  $\tau_t$  od parametrów systemu transmisyjnego, a ściślej od szybkości modulacji  $v$  i liczby bitów  $N$  w telegramie jest określona [2] następująco:

$$\tau_t = N \frac{1}{v}. \quad (4)$$

## 3. ZALEŻNOŚCI MIĘDZY PARAMETRAMI

Uwzględniając zależność czasu trwania telegramu od parametrów systemu APP — wyrażenie 2 i parametrów systemu transmisyjnego — wyrażenie 4, otrzymujemy równanie 5 uzależniające wybrane parametry obu systemów, tj. APP i systemu transmisyjnego.

$$\frac{S_T}{V_{max}} = \frac{N}{v}. \quad (5)$$

Przekształcając odpowiednio wyrażenie 5 można wyznaczyć:

— maksymalną liczbę bitów telegramu

$$N_{max} = \frac{S_T v}{V_{max}} \quad (6)$$

— minimalną szybkość modulacji

$$v = \frac{N V_{max}}{S_T} \quad (7)$$

— maksymalną prędkość pociągu

$$V_{max} = \frac{v S_T}{N} \quad (8)$$

w zależności od pozostałych parametrów systemu automatycznego prowadzenia pociągu — APP i systemu przekazywania informacji. Określony system auto-

matycznego prowadzenia pociągu przy założonych: drodze jaką przejeżdża pociąg przy odbiorze 1 telegramu  $S_T$  oraz prędkość maksymalną pociągu  $V_{max}$  charakteryzują dwa parametry: liczba wszystkich bitów telegramu  $N$  oraz wymagana szybkość modulacji  $v$ , która determinuje wybór systemu transmisyjnego.

Zatem z punktu widzenia systemu transmisyjnego, dowolny system automatycznego prowadzenia pociągu — APP charakteryzuje dwa parametry: całkowita liczba bitów w telegramie  $N$  i wymagana szybkość modulacji  $v$  (wyrażenie 7).

Traktując zagadnienie odwrotnie można dla określonego systemu przekazywania informacji, charakteryzowanego parametrem szybkość modulacji  $v$ , wyznaczyć liczbę  $N$  wszystkich bitów jaka przypada na jednostkę przejeżdżanej drogi  $S_T$ , w zależności od maksymalnej prędkości pociągu (wyrażenie 6).

Znając parametry systemu APP — liczba wszystkich bitów w telegramie  $N$  oraz systemu transmisyjnego — szybkość modulacji  $v$  przy pozostałych założeniach jak poprzednio można wyznaczyć maksymalną prędkość pociągu kursującego na linii z tymi systemami — wyrażenie 8.

## PODSUMOWANIE

W opracowaniu wykazano na ścisłe uzależnienie między niektórymi parametrami systemu automatycznego prowadzenia pociągu — APP i systemem przekazywania informacji z toru do pociągu.

## BIBLIOGRAFIA

1. H. Karbowski i inni: Teoria i struktury systemów automatycznego prowadzenia pociągu. CPBP 02.19.02.03. Praca niepublikowana. Politechnika Łódzka, Zakład Trakcji Elektrycznej, Łódź, 1988
2. Z. Baran i inni: Podstawy transmisji danych. Warszawa: WKŁ, 1982

A. KARBOWIAK

## RELATIONS BETWEEN THE PARAMETERS OF THE AUTOMATIC TRAIN DRIVING SYSTEM APP AND OF THE SYSTEM OF INFORMATION TRANSFER FROM THE TRAIN TRACK

### Summary

Intensive works on automatic train driving systems and on systems of information transfer from the track to the train, are under way. Attempts to introduce given parameters characterising both systems and to determine correlations between these parameters are presented. The parameter: way passed by the train when one telegram  $S_1$  characterising the automatic train driving system APP, is introduced. The relations between the parameter of the APP system and the parameters of the information transfer system: the modulation rate and the number of bits in the telegram, are derived.

537.72:534.213.4

# Defektoskopia wiropądowa rur nieferromagnetycznych spawanych

ANNA LEWIŃSKA-ROMICKA

*Centrum Uczelniano-Przemysłowe Metrologii i Systemów Pomiarowych, Politechnika Warszawska*

*Otrzymano 1989.07.30*

*Autoryzowano do druku 1989.08.03*

W artykule opisano defektoskop wiropądowy, przeznaczony do wykrywania wad złączy spawanych rur miedzianych i mosiężnych podczas procesu ich wytwarzania. Przedstawiono przykładowe wady złączy rur. Przedyskutowano zagadnienie obiektów wzorcowych.

## 1. WPROWADZENIE

Opisano zagadnienia, jakie należy wziąć pod uwagę przy rozpatrywaniu i rozwiązywaniu problemów nieniszczących badań materiałów, w szczególności problemów wykrywania wad obiektów w przypadkach, gdy stosowalna jest metoda prądów wirowych.

Urządzenia wiropądowe do kontroli obiektów, podobnie jak i inne urządzenia do badań nieniszczących, mogą być rozpatrywane jako urzeczywistnienie systemów pomiarowych, które spełniają szereg ogólnych funkcji. Najważniejsze z nich są następujące:

1. Wprowadzenie i propagacja energii w obiekcie badanym, która to energia jest w pewien sposób zmieniana pod wpływem obiektu badanego,
2. Detekcja tej energii,
3. Interpretacja dekodowanej energii,
4. Prezentacja informacji w formie dekodowanej energii — operatorowi lub wykorzystanie informacji do sterowania urządzeniami wykonawczymi,
5. Podawanie obiektów w sposób zgodny z procesem ich wytwarzania oraz/lub z wymaganiami kontroli.

Decyzja odnośnie postaci urządzenia kontrolnego, jakie ma być realizowane lub zastosowane dla spełnienia powyższych funkcji, w odniesieniu do konkretnego problemu kontroli, wymaga wzięcia pod uwagę szeregu czynników. Zostaną one przedstawione w postaci zestawu pytań. Dalsze wywody niniejszego artykułu nawiązywać będą do znacznej części spośród tych pytań.

1. Jaki jest cel przeprowadzania badania nieniszczącego?

2. Jaki jest kształt obiektu badanego, jakie są jego rozmiary i z jakiego materiału jest wykonany?
3. Jaki rodzaj wad ma być wykrywany i określany?
4. Czy jest istotne rozróżnianie poszczególnych wad od siebie?
5. Czy problem polega na detekcji wad, czy też na pomiarze ich wielkości?
6. Jaka jest najmniejsza istotna wada?
7. Jaka jest wielkość produkcji, która ma być kontrolowana?
8. Jaka jest technologia wytwarzania obiektu?
9. W jaki sposób i kiedy pojawiają się wady obiektu?
10. Czy badany jest materiał i/lub wyrób gotowy?
11. W jakim stadium wytwarzania obiektu badanie ma być zastosowane?
12. Czy ma być realizowana kontrola „on-line” czy „off-line”?
13. Czy docelowa powierzchnia obiektu jest już osiągnięta, czy też będzie miała zastosowanie dodatkowa obróbka?
14. Jaki jest stan: metalurgiczny i chropowatość powierzchni materiału lub obiektu?
15. Czy próbki wybrane do badań są typowe dla danej produkcji?
16. Czy mogą być uzyskane obiekty wzorcowe?
17. Czy mogą być wytworzone odpowiednie wzorcowe wady sztuczne?
18. Jakie (uzupełniające) badania niszczące mogą być zastosowane?
19. Jaki rodzaj informacji odnośnie obiektu ma być uzyskany?

Lista przytoczonych, najbardziej istotnych, wzajemnie się uzupełniających pytań wskazuje typy zagadnień, które są istotne przy konstrukcji i wykorzystaniu aparatury wiropądowej. W odniesieniu do przedstawionej wyżej listy funkcji, jakie powinno spełniać urządzenie kontrolne, najistotniejszym zagadnieniem jest wprowadzenie i propagacja energii w obiekcie. Metoda prądów wirowych może mieć zastosowanie jedynie do obiektów przewodzących prąd elektryczny. Jeśli problem dotyczy wykrywania wad powierzchniowych, to metoda prądów wirowych świetnie się nadaje do tego celu. Jeśli przy tym ostateczna postać powierzchni obiektu badanego nie została osiągnięta na etapie, w którym przeprowadzona ma być kontrola, to problem powinien być rozpatrywany jako zagadnienie detekcji wad podpowierzchniowych lub jako wykrywanie wad powierzchniowych w odpowiednim etapie wytwarzania. W tym przypadku wyniki kontroli wiropądowej nie mogą jednak stanowić o końcowym dopuszczeniu obiektów.

Następnym pod względem ważności zagadnieniem jest dobór typu przetwornika wiropądowego, odpowiedniego do zamierzonego celu i technika użycia tego przetwornika. Jest to problem bardzo obszerny i w tym miejscu podanych będzie tylko kilka uwag. Informacja, jaka uzyskiwana jest w urządzeniu kontrolnym, pochodzi jedynie z przetwornika pomiarowego. Jeśli problem kontroli wymaga wyszukanego oprzyrządowania (m.in. przetwornika), to kontrola może mieć ograniczoną prędkość i/lub urządzenie będzie złożone i kosztowne. Dla przypadku wykrywania nieciągłości dobór przetwornika, który ma postać zespołu cewek indukcyjnych, następuje w zależności od m.in. kształtu, rozmiarów, rodzaju materiału obiektu badanego oraz konfiguracji przewidywanych wad. Problem odpowiednio komplikuje się m.in., jeśli równocześnie mają być wykrywane zarówno wady małe, jak i duże, bądź, gdy

konieczna jest ich lokalizacja lub, gdy ich wielkość ma być mierzona lub też, gdy poszczególne rodzaje wad powinny być rozróżniane. Czułości wykrywania wad zależy od stanu powierzchni, w tym jej chropowatości, odchylen od prostoliniowości obiektu, niejednorodności właściwości elektrycznych i magnetycznych materiału obiektu badanego oraz geometrii i stałości rozmiarów obiektu. Czynniki te są przyczyną występowania zakłóceń, które mogą w sposób istotny ograniczać czułość urządzenia w odniesieniu do wad.

Odrębnym zagadnieniem jest dobór częstotliwości pracy przetwornika. Na wybór tej częstotliwości mają wpływ następujące czynniki: rodzaj, kształt i rozmiary obiektu, rodzaj materiału obiektu, rodzaj i położenie wad, jakie mają być wykrywane oraz prędkość kontroli. Odpowiedni wybór częstotliwości pracy przetwornika stwarza przesłanki wykrywania wad obiektu na tle wpływu czynników zakłócających.

Detekcja sygnałów, w których zawarta jest informacja o ewentualnych wadach obiektu, realizowana jest w zależności od informacji wymaganej od urządzenia. Wybór układów detekcji sygnałów powinien być tak dokonany, aby można było odróżnić informacje istotne od nieistotnych i klasyfikować wady w zależności od ich typów i wielkości, przy czym może być tu uzyskany różny stopień inteligencji układu.

Prezentacja informacji, co stanowi czwartą funkcję urządzenia kontrolnego, zależy od sukcesu w detekcji i interpretacji sygnałów, uzyskiwanych z przetwornika.

Rozwiązanie problemu piątego — podawania, czy też prowadzenia obiektów przez przetwornik, wymaga wzięcia pod uwagę geometrii i masy poszczególnych obiektów oraz wielkości produkcji, a w przypadku kontroli potokowej — również prędkości wytwarzania. Przy konstruowaniu aparatury wiropądowej, z przeznaczeniem do badania określonych obiektów, wszystkie wymienione czynniki powinny być rozpatrzone zarówno wspólnie, jak i każdy indywidualnie.

Każdy problem kontroli powinien być rozpatrywany jako zagadnienie unikatowe ze względu na to, że istnieje możliwość konstrukcji urządzenia optymalnego. Konstruowanie uniwersalnych defektoskopów wiropądowych wydaje się być bowiem podejściem przestarzałym, choć możliwe jest użycie przetworników o różnych konfiguracjach oraz „kompletna” detekcja sygnałów, uzyskiwanych z przetworników.

Celowe jest przeprowadzanie badań nieniszczących obiektów u producentów, niemniej badania u odbiorców, choć z pozoru wydają się być nadmiarowymi, mogą się okazać pożytecznymi w wielu przypadkach. W wyniku przeprowadzenia badania u użytkownika mogą być wykryte wady pominięte u producenta, szczególnie, gdy użytkownik dysponuje inaczej zaprojektowanym urządzeniem kontrolnym, o innych nastawach oraz gdy dodatkowo przyjmowany jest odmienny standard wad.

W niniejszej pracy ograniczono się do prezentacji zagadnień, związanych z opracowaniem defektoskopu do potokowej kontroli jakości rur — przy wykorzystaniu przelotowego przetwornika wiropądowego. Załączono obszerny wykaz literatury przedmiotu [4—49, 53, 54, 67, 60—77, 79—83].

## 2. ZASTOSOWANIA DEFEKTOSKOPÓW WIROPRAĐOWYCH DO KONTROLI RUR

Defektoskopy wiroprowadowe mogą być opracowane z przeznaczeniem do kontroli jakości rur, otrzymywanych metodą zwiłania z taśmy i wzdłużnego ich spawania. Taki cel kontroli założono przy opracowaniu niżej opisanego defektoskopu wiroprowadowego. Rury te, po spawaniu, stanowią produkt gotowy, bądź poddawane są dalszej obróbce, jak przeciąganie, kalibrowanie, nakładanie powłok galwanicznych lub plastycznych i innych. Zwykle wymagana jest:

- wysoka tolerancja rozmiarów rur,
- szczelność spoin,
- wysoka wytrzymałość mechaniczna rur.

W zależności od przyjętego procesu technologicznego defektoskopy wiroprowadowe mogą znaleźć zastosowanie

- dla rur spawanych (nie poddawanych przeciąganiu) jako urządzenia kontrolujące proces spawania oraz równocześnie jako urządzenia sortujące gotowe produkty,

- dla rur spawanych, a następnie przeciąganych (zwykle na oddzielnych liniach technologicznych) jako niezależne urządzenia, oddzielne — ustawione w linii spawania i kontrolujące proces spawania, oddzielne — w linii przeciągania i służące do badania oraz sortowania gotowych produktów.

Defektoskopy wiroprowadowe mogą znaleźć zastosowanie przy badaniu i sortowaniu:

- rur wytwarzanych dowolną technologią (przeciągania, spiralnego lub wzdłużnego zwiłania, a następnie spawania lub zgrzewania), wykonanych z metali kolorowych (miedź, mosiądz, brąz, aluminium) lub metali żelaznych (stale wysokostopowe i konstrukcyjne),

- prętów i drutów wykonanych z dowolnych metali oraz przy badaniu rur zainstalowanych

- rurociągów,

- rur w wymiennikach ciepła — w szczególności w energetyce jądrowej.

Jakość rur spawanych (zwiłanych z taśmy) zależy głównie od jakości złącz. Rury z metali kolorowych spawane są metodą TIG (tungsten inert gas welding; elektroda wykonana jest z wolframu, osłonę stanowi gaz obojętny — argon). Złącza przebiegają wzdłuż długości rur [56].

Typowymi zastosowaniami rur, wykonanych z miedzi i mosiądzu, są wymienniki ciepła, osprzęt spawalniczy, instalacje, otuliny części grzejnych, armatura i osprzęt oświetleniowy, przy czym jako bardzo odpowiedzialne należy wymienić zastosowania np. w wymiennikach ciepła.

Badane rury spawane stanowią zarówno produkt finalny, jak i półprodukt do obróbki plastycznej na zimno. W procesie ciągnięcia otrzymywane są z tych rur rury o mniejszych średnicach, również kapilarne. Wady złącz rur spawanych jak np. częściowy przetop, są przyczyną powstawania nieszczelności rur otrzymywanych w procesie dalszego ciągnięcia.

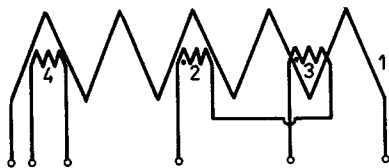
Rury poddane badaniom były miedziane i mosiężne o przekroju okrągłym, wykonane z następujących gatunków materiałów: M1E, M1R, M2G i M63. W przyszłości przewidywane jest podjęcie zagadnień badania rur z brązów i ew. aluminium.

Średnice i grubości ścianki rur wynosiły odpowiednio 6,5 mm (0,5 mm; 0,6 mm); 10,0 mm (0,5 mm; 0,6 mm) i 16,0 mm (0,5 mm; 0,6 mm; 1,0 mm) oraz 18,0 mm (0,6 mm; 1,0 mm; 1,2 mm). Rury spawane są z prędkością 3—8 m/min. Przed wprowadzeniem do przetwornika wiroprowdowego są odpowiednio schładzane. Rury wytwarzane są w odcinkach o długości od 3 m do 6 m lub w kręgach o długości od 20 m do 50 m. Kontrolę podlegają rury wytwarzane w jednej linii technologicznej.

Opracowany defektoskop wiroprowdowy przeznaczony jest do potokowej kontroli rur — na obecność wad złącz-bezpośrednio po procesie spawania i następującego po nim wyżarzania normalizującego (to ostatnie — tylko w odniesieniu do rur mosiężnych). Uzyskane sygnały wykorzystywane są m.in. do sortowania rur gotowych. Wykrywane wady złącz rur klasyfikowane są na trzy grupy: małe, średnie oraz duże. Te kategorie wad to, zgodnie z wymaganiami użytkownika, odpowiednio 10%—25% brak przetopu w stosunku do grubości ścianki rur, brak spoiny lub dziura na długości nie większej niż 2 mm oraz brak spoiny lub dziura na długości większej niż 2 mm. Sygnały otrzymywane przy występowaniu poszczególnych kategorii wad podawane są do odpowiednich trzech wyjść defektoskopu. Z defektoskopem współpracują następujące urządzenia badawcze, decyzyjne, rejestrujące i wykonawcze: sterownik mikroprocesorowy wraz z odpowiednim impulsatorem, znakownik, piła latająca wraz z układem sortującym, sygnalizatory optyczny i akustyczny oraz rejestrator statystyczny wad. Wymienione urządzenia wbudowano do istniejącej linii spawania rur w Hucie Będzin.

### 3. PRZETWORNIK WIROPROWDOWY

Do kontroli jakości złącz rur spawanych opracowano zewnętrzny przelotowy przetwornik generacyjny złożony — pracujący w układzie stosowanym w metodzie bezwzględnej i różnicowej. Badane rury przesuwane są poprzez przetwornik. Konfigurację uzwojeń przetwornika przedstawiono na rys. 1. Sekcja bezwzględna przetwornika umożliwia uzyskanie sygnału elektrycznego, trwającego przez czas przechodzenia wady przez przetwornik, natomiast sekcja różnicowa umożliwia uzyskanie



Rys. 1. Schemat elektryczny przetwornika w układzie bezwzględnym i różnicowym; 1 — uzwojenie wzbudzające, 2—3 — uzwojenia wyjściowe przetwornika różnicowego, 4 — uzwojenie wyjściowe przetwornika bezwzględnego

różnoimiennych sygnałów dla początku i końca wady. Sygnały sekcji bezwzględnej przetwornika wiroprowadowego wykorzystywane są do wykrywania wad „dużych” przelotowych (brak szwu na znacznej długości, duże dziury, powstałe wskutek wypaleń elektrodą) i dla sygnałów pochodzących od tych wad przewidziano priorytet podczas procesu ich obróbki w układach elektronicznych defektoskopu. Sygnały sekcji różnicowej przetwornika wiroprowadowego wykorzystywane są do wykrywania wad „małych” i „średnich” odpowiednio — przelotowych takich, jak pęknięcia i brak przetopu i wad nieprzelotowych, takich jak przetop częściowy. Należy podkreślić, że opracowany przetwornik wiroprowadowy przelotowy, z zasady swojego działania służy do wykrywania wad wzdłużnych, a takie wady występują w złączach badanych rur oraz, że przetwornik taki odznacza się właściwością „sumowania wad”, które ma miejsce np. w przypadku wystąpienia w danym przekroju poprzecznym rury kilku wad jednocześnie (np. pęknięcie i porowatość złącza lub odmienny od prawidłowego kształt złącza, względnie przemieszczenie spawanych krawędzi taśmy i pęknięcie złącza, zmiana grubości ścianki i częściowy przetop, zmiana przewodności materiału rury i przepalenie itd.). Dzięki zastosowaniu przetwornika w układzie różnicowym możliwe jest zmniejszenie niektórych wpływów tj. zmian właściwości fizykochemicznych lub rozmiarów, gdyż sygnał takiego przetwornika zależy od różnicy stanu obiektu kontrolowanego w miejscach odpowiadających położeniu odpowiednich uzwojeń.

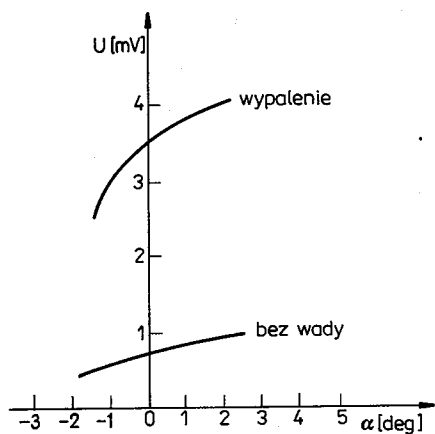
Jak wynika z przeprowadzonych badań, sygnał przetwornika wiroprowadowego zależy nie tylko od występowania wad ale i od niejednorodności struktury i zmian rozmiarów rur badanych, a także ich nieosiowego i niecentrycznego prowadzenia przez przetwornik. Sygnały przetwornika wiroprowadowego w układzie bezwzględnym, opracowanego do kontroli rur, analizowane na płaszczyźnie zmiennej zespolonej, przedstawiono w [37]. Jak wynika z podanych tam rozważań i pomiarów, sygnał przetwornika wiroprowadowego pozostaje w silnej zależności od wymienionych wyżej czynników zakłócających wykrywanie wad. W celu zmniejszenia (ew. wyeliminowania) wpływu tych czynników, konieczne jest zastosowanie analizy fazoczułej. Jednakże, wady nieprzelotowe i przelotowe rur wywołują sygnały o odmiennych fazach (dla mniejszych częstotliwości pracy przetwornika). Ograniczenia aparaturowe uniemożliwiały prześledzenie tych zależności dla częstotliwości większych od 10 kHz. Przy zastosowaniu jednak odpowiednio wyższej częstotliwości pracy defektoskopu faza napięcia przetwornika wywołanego wadami przelotowymi i nieprzelotowymi będzie miała taki sam znak, choć kąt przesunięcia fazowego będzie przybierał różne wartości. Stanowi to jedną z przesłanek realizacji, koniecznej tu, fazoczułej obróbki sygnałów. Obróbka fazoczuła została zastosowana do analizy sygnałów obu sekcji przetwornika.

Niżej przedstawiono wpływ nieosiowego prowadzenia rury — jako przykład jednego z wpływów zakłócających wykrywanie wad przetwornikami wiroprowadowymi przelotowymi. Na rys. 2 przedstawiono napięcie przetwornika w układzie różnicowym w zależności od wartości kąta przekoszenia rury w stosunku do osiowego, centrycznego prowadzenia, dla dwóch przypadków: dla rury bez wady i rury z dziurą, powstałą wskutek wypalenia elektrodą. Pomiar przeprowadzono dla rury mosiężnej średnicy

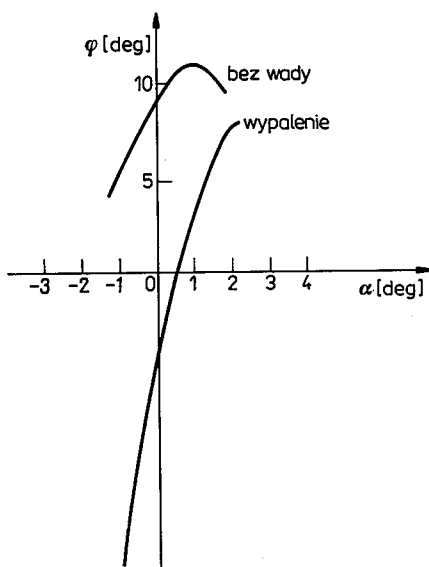
16 mm i grubości ścianki 1,0 mm. Częstotliwość pracy przetwornika wynosiła 9 kHz.

Jak widać z rys. 2, nieprawidłowe prowadzenie rury powoduje zmiany napięcia wyjściowego przetwornika, co jest szczególnie widoczne dla rury z wadą.

Na rys. 3 przedstawiono, dla tej samej rury i częstotliwości pracy, kąt fazowy  $\varphi$  napięcia przetwornika przy nieosiowym prowadzeniu rury. Obserwuje się stosunkowo



Rys. 2. Napięcie przetwornika różnicowego przy nieosiowym prowadzeniu rury



Rys. 3. Kąt fazowy napięcia przetwornika różnicowego przy nieosiowym prowadzeniu rury

szybką zmianę tego kąta w zależności od kąta  $\alpha$  przekoszenia rury badanej, szczególnie dla rury z wypaleniem.

Problemy te skłoniły autorów\* defektoskopu do takiego wyboru zależności fazowych w układach defektoskopu, aby w torze obróbki sygnałów uzyskiwanych z przetwornika bezwzględnie eliminować wpływ zmian rur, a w torze obróbki sygnałów uzyskiwanych z przetwornika różnicowego — eliminować wpływ nieprawidłowego prowadzenia rur.

#### 4. CZĘSTOTLIWOŚĆ PRACY DEFEXTOSKOPU

W pracach [35, 37] przedstawiono szczegółowe rozważania odnośnie częstotliwości pracy przetworników, opracowanych do kontroli złączy rur spawanych. M.in. w tabeli 3 pracy [37] zestawiono, obliczone wg kryterium dla rur cienkościennych, tzw. częstotliwości graniczne dla rur miedzianych i mosiężnych (spawanych) wytwarzanych

\* A. Lewińska-Romicka, S. Piskorski

w Polsce. Częstotliwości te, w zależności od materiału, z których wykonywane są rury, ich średnicy zewnętrznej i grubości ścianki, zawierają się w zakresie od 459 Hz do 12281 Hz.

W celu wyboru częstotliwości pracy prototypu defektoskopu uwzględniono następujące przesłanki:

- 1) materiały, z których wykonywane są rury badane,
- 2) zakres średnic zewnętrznych badanych rur,
- 3) grubość ścianek badanych rur.

Można dobrać optymalną częstotliwość pomiarową dla przypadku kontroli rur o poszczególnych średnicach zewnętrznych, określonych grubościami, wykonanych z różnych materiałów. Jednak w związku z zastosowaniem opracowanych przetworników w prototypie defektoskopu do kontroli rur, wykonanych z materiałów nieferromagnetycznych, celowy wydawał się wybór jednej częstotliwości pracy. Zagadnienie to nie jest krytyczne, gdyż, jak to wynika z [17], składowa zmiany napięcia wyjściowego przetwornika, wywołana np. przez nieciągłość, prostopadła do „kierunku” zmian średnicy zewnętrznej, przy wartościach stosunku częstotliwości pracy do częstotliwości granicznej w zakresie około 0,8—8 zmienia się stosunkowo niewiele. Ponadto wzięto pod uwagę okoliczność, że początkowo defektoskop będzie wykorzystywany do kontroli rur tylko średnicy 16 mm.

Częstotliwości graniczne dla rur miedzianych średnicy 16 mm i grubości ścianki w zakresie od 0,5 mm do 1,0 mm wynoszą od 1145 Hz do 613 Hz, a dla rur mosiężnych — od 4503 Hz do 2412 Hz. Przyjęto założenie, że stosunek częstotliwości pracy defektoskopu do częstotliwości granicznych dla badanych rur powinien zawierać się w zakresie od około 3 do 25. W ten sposób określono częstotliwość pracy prototypu defektoskopu jako równą 15 kHz.

W przypadku innych niż założony cel kontroli, na wybór częstotliwości pracy wpływa, jak zaznaczono wyżej, również prędkość przesuwu obiektu badanego. Jeśli przyjąć np. założenie, że wada o długości 1 mm, występująca w dowolnym obiekcie, przesuwać się powinna poprzez uzwojenie wyjściowe przetwornika przez czas odpowiadający conajmniej dziesięciu okresom napięcia zasilającego przetwornik, ze względu na stany nieustalone, to częstotliwość pracy  $f \geq 166,7 A$ , przy czym  $A$  — jest prędkością przesuwu obiektu badanego przez przetwornik, wyrażoną w metrach na minutę. Częstotliwość pracy  $f$  wyrażona będzie w hercach. Należy nadmienić, że w przypadku potokowej kontroli rur stalowych zgrzewanych i prętów walcowanych oraz drutów ciągnionych, prędkości przesuwu wynoszą odpowiednio około 50 m/s oraz 50 m/s — 120 m/s [7, 76].

Im większa jest prędkość przesuwu obiektu podczas badań, tym większa musi być częstotliwość pracy przetwornika, co z kolei wpływa na zmniejszenie głębokości penetracji prądów wirowych w głąb ścianki rur i może uniemożliwić wykrywanie wad zlokalizowanych od strony wewnętrznej ścianki rur.

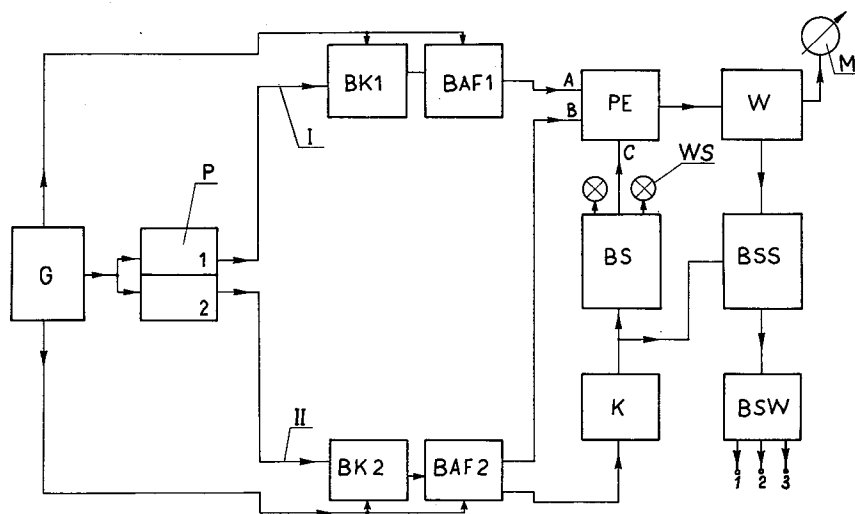
## 5. OPIS DEFJEKTOSKOPU

Schemat blokowy układu defektoskopu, model DW-31, przedstawiono na rys. 4 [42].

Sygnal z generatora G doprowadzany jest do sondy pomiarowej P oraz do obu bloków analizy fazoczułej BAF1 i BAF2, gdzie jest wykorzystywany zarówno w procesie kompensacji, jak i w procedurze wykrywania wad. Zasadniczym zadaniem generatora jest:

- ysterowanie przetwornika pomiarowego w celu wzbudzenia w badanym obiekcie prądów wirowych o określonej częstotliwości,
- ustalenie amplitudy i fazy napięcia do kompensacji,
- synchronizacja procesu analizy fazoczułej w procedurze wykrywania wad.

Blok analizy fazoczułej składa się z dwóch części BAF 1 i BAF 2: jedna z nich współpracuje z przetwornikiem bezwzględny, natomiast druga — z przetwornikiem różnicowy. Oba przetworniki zawarte są w sondzie pomiarowej P, stanowiącej podstawowy podzespół defektoskopu. Zastosowanie detekcji fazoczułej umożliwia znaczne zmniejszenie zakłóceń wywołanych zmianami średnicy, przekoszeniem i wibracjami rury spawanej, w linii technologicznej.



Rys. 4. Schemat blokowy defektoskopu wiroprądowego, model DW-31

Blok sterowania BS steruje komutacją w blokach analizy fazoczułej w procedurze kompensacji i selekcją sygnałów wyjściowych w procedurze wykrywania wad.

Blok rejestracji i sygnalizacji wad BSS nadaje ostateczną postać sygnałom niosącym informację o wadliwości badanej rury. Z defektoskopu uzyskujemy 3 sygnały wyjściowe. W każdym z nich zawarta jest informacja o istnieniu określonego typu wad odpowiednio: małych, średnich i dużych. Obecność np. wady małej w badanej rurze zauważana jest przez urządzenie wtedy, gdy po procesie analizy sygnałów z przetwor-

nika wartość napięcia doprowadzonego do wskaźnika analogowego przekracza ustawioną przez użytkownika w procesie kalibracji dolną wartość progową dla wad małych. Taka sama sytuacja występuje dla wad średnich i dużych. Pojawienie się wady sygnalizuje świecenie się lampki kontrolnej odpowiedniego rodzaju wady. Wielkość wykrywanych wad pokazywana jest na analogowym wskaźniku odczytowym przyrządu. Skala wskaźnika jest w przybliżeniu proporcjonalna do wielkości wad.

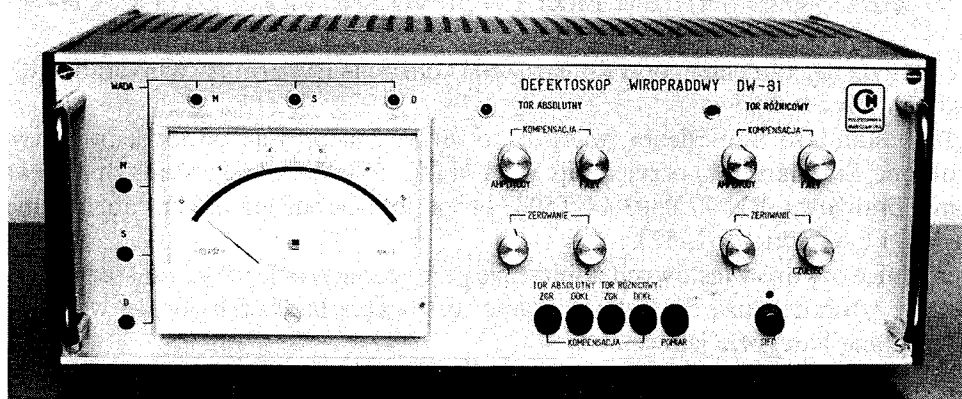
Sygnały wyjściowe defektoskopu dostosowane są do współpracy z systemem mikroprocesorowym, zastosowanym w linii technologicznej. Mają one postać impulsów TTL, gdzie wartość logiczna „1” odpowiada pojawieniu się wady w badanym obiekcie. Niżej zostanie opisany układ defektoskopu w sposób bardziej szczegółowy [42, 44].

Generator G zasila dwusekcyjny przetwornik pomiarowy wiroprowodowy P o sekcjach 1,2, przez który przesuwane są badane rury. Sekcja 1 przetwornika P została zrealizowana w wersji różnicowej, tj. takiej, w której sygnał wyjściowy zależy od różnicy stanu obiektu w dwóch sąsiadujących miejscach, znajdujących się w danej chwili w obszarze działania dwóch krótkich symetrycznych uzwojeń, połączonych przeciwsośnie. Uzyskano w ten sposób dużą czułość wykonania wad przy stosunkowo niewielkich wymiarach uzwojeń sekcji. Jednocześnie przy takiej konfiguracji tłumione są niektóre wpływy zakłócające pomiar, którymi są tu zmiany średnicy i grubości ścianki rur oraz zmiany struktury i składu chemicznego materiału rur, o ile zmiany te następują na odcinku dłuższym niż obszar uzwojeń sekcji różnicowej przetwornika.

Sekcja 2 przetwornika P została zrealizowana w wersji bezwzględnej, tj. takiej, w której sygnał wyjściowy zależy od stanu obiektu w obszarze działania odpowiedniego uzwojenia. Uzyskano w ten sposób możliwość wykrywania wad o stałej głębokości, położonych na całej długości obiektów badanych. Sekcja bezwzględna przetwornika umożliwia wykrywanie wad znacznych co do wielkości. Na wyjściu poszczególnych sekcji przetwornika uzyskiwane jest napięcie o amplitudzie i fazie zależnej od stanu odcinka badanej rury, znajdującego się w danej chwili w obszarze uzwojeń przetwornika.

Sygnały wyjściowe sekcji 1 i 2 przetwornika P sterują dwoma kanałami analizy sygnałów I i II. Kanały I i II zawierają bloki analizy fazowej BAF1 i BAF2 sygnałów sekcji 1, 2 przetwornika P. Dzięki temu możliwe jest wyodrębnienie napięć o fazie wywołanej pojawieniem się wady w rurach oraz stłumienie wpływów zakłócających, tj. zmian średnicy i grubości ścianki rur oraz nieosiowego ich prowadzenia poprzez przetwornik P. Sygnały do bloków analizy fazowej BAF1 i BAF2 doprowadzane są z wyjść sekcji 1 i 2 przetwornika P poprzez bloki kompensacji BK1 i BK2. Bloki kompensacji BK1 i BK2 oraz analizy fazowej BAF1 i BAF2 sterowane są napięciem odniesienia, pobieranym z generatora G. Kanał I umożliwia analizę sygnałów, wywołanych przez wady „małe” i „średnie”, a kanał II — przez wady duże.

Napięcia z wyjść bloków analizy fazowej BAF1 i BAF2 doprowadzane są do wejść A i B przełącznika elektronicznego PE. Do wejścia C przełącznika doprowadzane jest napięcie z bloku analizy fazowej BAF2 kanału II obróbki sygnałów z sekcji bezwzględnej 2 przetwornika P, przy czym napięcie to doprowadzane jest poprzez komparator K oraz blok sterowania BS. W komparatorze K zrealizowano płynną regulację czułości defektoskopu na wady „duże” — na podstawie sygnału sekcji



Fot. 5. Defektoskop wiroprowdowy, model DW-31

bezwzględnej 2 przetwornika P, tak, że napięcie przekraczające wybrany próg czułości podawane są do bloku selekcji sygnałów BSS.

Do wyjścia bloku sterowania BS podłączono wskaźniki świetlne WS dla wskazywania, który kanał analizy w danej chwili pracuje.

Sygnał z wyjścia przełącznika elektronicznego PE podawany jest do wzmacniacza prądu stałego W. Sygnały z wyjścia wzmacniacza W doprowadzane są do miernika odchyłowego M oraz do bloku selekcji sygnałów BSS. Wskazanie miernika odchyłowego M odpowiada stanowi odcinka obiektu znajdującego się w przetworniku P. W bloku selekcji sygnałów BSS następuje selekcja sygnałów, uzyskiwanych z kanału I analizy sygnałów, otrzymywanych z sekcji różnicowej 1 przetwornika, tzn. realizowana jest selekcja sygnałów, odpowiadających wadom, na kategorie „małe” i „średnie”.

Do bloku selekcji sygnałów BSS podawane jest również napięcie z wyjścia komparatora K, w którym to napięciu zawarta jest informacja w odniesieniu do wad „dużych”, otrzymywana na podstawie analizy w kanale II sygnałów uzyskiwanych z wyjścia sekcji bezwzględnej 2 przetwornika P.

Sygnały z wyjścia bloku selekcji BSS doprowadzone są do układu sygnalizacji wad BSW. W bloku tym zrealizowano świetlną sygnalizację wymienionych trzech kategorii wad. Blok sygnalizacji wad BSW wyposażono w trzy wyjścia 1, 2, 3, do których doprowadzone są napięcia odpowiadające wadom „małym”, „średnim” i „dużym”. Napięcia z wyjść bloku sygnalizacji wad BSW służą do sterowania odpowiednimi urządzeniami wykonawczymi, takimi jak np. piła do wycinania wadliwych odcinków rur. Napięcia te mogą być doprowadzane do układu mikroprocesorowego rejestracji stanu poszczególnych odcinków obiektów w zależności od wad. Na fot. 5 przedstawiono defektoskop wiroprowdowy, model DW-31.

## 6. RURY TESTOWE DLA DEFEKTOSKOPII WIROPRAĐOWEJ Z ZASTOSOWANIEM PRZETWORNIKÓW PRZELOTOWYCH

Istotnym zagadnieniem przy konstrukcji i aplikacjach aparatury wiroprowadowej jest sprawa wzorców wad.

Jako standard odniesienia (wady testowe) można przyjąć odpowiednie otwory obrobione. Standard taki przyjęto np. w RFN [47, 55]. W Polsce została opracowana norma branżowa BN-77/0803-14 [50], która jest odpowiednikiem norm amerykańskiej i brytyjskiej [51, 52].

Średnice otworów testowych dobiera się przeważnie w zależności od średnicy rury. Te wady sztuczne służą do wyboru granic sortowania badanych rur, ze względu na poszczególne kategorie jakościowe.

Należy podkreślić, że wielkości testowych wad sztucznych nie mogą służyć bezpośrednio jako miara wad naturalnych, jakie zostaną zarejestrowane przy określonej czułości urządzenia kontrolnego. Wynika to przede wszystkim z geometrii wad naturalnych, które najczęściej mają nieregularny kształt. W szczególności uwaga ta dotyczy wad złącz rur, do wykrywania których przeznaczony jest opisywany defektoskop wiroprowadowy, model DW-31.

Określenie rodzajów i wielkości wad złącz rur, wykrywanych przy danej czułości, wymaga przeprowadzenia badań defektoskopem wiroprowadowym, uzupełnionych następnie niszczącymi badaniami metalograficznymi. Rury testowe ze wzorcowymi wadami sztucznymi mogą być łatwo wytworzone u użytkownika aparatury, a więc nie występuje konieczność uzyskiwania tych wzorców w daleko położonej „centrali”. Możliwe są tu dwa sposoby podejścia. Pierwszy z nich polega na tym, że określone otwory testowe wykonane zostaną w rurze o określonej średnicy, drugi — na tym, że otwory te wprowadzane są w rurach o poszczególnych średnicach [55]. Otwory testowe wykonywane są w rurach będących w takim samym stanie, co badana partia rur [50]. Rury te wybiera się spośród tych, dla których badania defektoskopem wykazały najmniejsze wartości amplitudy zarejestrowanych sygnałów, przy czułości większej niż przewidywana czułość badania.

Norma zachodnio-niemiecka [55] podaje zestaw 50 wad testowych. Podano je dla rur o średnicy zewnętrznej 25,0 mm i grubości ścianki 2,0 mm. Do wyboru są cztery określone materiały rur testowych, a mianowicie miedź, mosiądz, stal austenityczna (nieferromagnetyczna) oraz stal ferromagnetyczna. Wady testowe wykonywane są w pięciu rurach o długości 2 m każda. Wady testowe różnią się pod względem typu, rozmiarów i położenia na rurze. Można je zebrać w siedmiu grupach:

1) rowki prostokątne wzdłużne o długości 30 mm i szerokości 0,5 mm, o zróżnicowanej głębokości (20%, 40% i 60% grubości ścianki rury), od strony zewnętrznej ścianki rur

2) rowki pierścieniowe o szerokości 0,5 mm i o różnej głębokości (20% i 40% grubości ścianki rury), od strony wewnętrznej i zewnętrznej ścianki rur

3) kanałki frezowane podłużne o różnej długości (od 2,5 mm do 40 mm), od strony zewnętrznej ścianki rur

4) grupy trzech otworów średnicy 1 mm, w różnych odstępach (od 4 mm do 20 mm)

5) otwory średnicy 1 mm, w stałych odstępach i położone co  $120^\circ$  w kierunku obwodowym

6) otwór o podwójnej średnicy (tj. 2 mm), umieszczony w odległości 100 mm od początku rury

7) otwory średnicy 1 mm, przy końcu wszystkich pięciu rur.

Ten bogaty asortyment wad sztucznych umożliwia badania aparatury odnośnie czułości w odniesieniu do wad krótkich i długich, odpowiednio zewnętrznych i wewnętrznych, badanie wpływu dokładności prowadzenia oraz badanie wpływu końców na wykrywalność wad rur. W metodzie wiroprowdowej z zastosowaniem przetworników przelotowych na końcu badanych obiektów istnieje bowiem zawsze strefa niekontrolowana.

Polska norma branżowa [50] jako otwory testowe przewiduje otwory wywiercone na wylot przez jedną ściankę rury, prostopadle do jej osi, położone w odległości 500 mm od siebie i co  $120^\circ$  w kierunku obwodowym rury. Cytowana norma dotyczy rur bez szwu, z przeznaczeniem do wymienników ciepła oraz rur ogólnego przeznaczenia. I tak np. średnice otworów testowych dla rur, przeznaczonych do wymienników ciepła, o średnicach w zakresie od 6,0 mm do 20,0 mm wynoszą  $(0,6 + 0,05)$  mm, a dla rur ogólnego przeznaczenia o średnicach w zakresie od 6,0 mm do 10,0 mm —  $(0,75 + 0,05)$  mm i dla rur o średnicach 10,0 mm do 20,0 mm —  $(0,90 + 0,05)$  mm.

Rozmiary wszystkich wymienionych wyżej wad testowych świadczą pośrednio o (ograniczonej) czułości metody wiroprowdowej z zastosowaniem przetworników przelotowych. Sygnał przetwornika wiroprowdowego jest bowiem ogólną miarą stanu odcinka obiektu, a więc zależy od parametrów fizykochemicznych, rozmiarów obiektu i ew. występowania w nim wad.

Należy podkreślić, że do chwili obecnej brak jest normy dotyczącej wymagań jakościowych odnośnie spawanych rur z miedzi i mosiądzu, szczególnie w kontekście badań metodą indukcyjną. Badania defektoskopowe rur ze szwem stanowią znacznie trudniejszy problem niż rur bez szwu. I tak np. w badanych rurach niemagnetycznych, ze szwem, obserwuje się znaczne nieregularności geometryczne, wynikające z procesu formowania rury szczelinowej oraz nieregularności właściwości fizykochemicznych, wynikające z procesu spawania i wyżarzania. Wymienione czynniki zakłócają proces wykrywania wad złącz rur metodą wiroprowdową.

## 7. PRZYKŁADOWE WYNIKI BADAŃ RUR Z WADAMI TESTOWYMI

W celu określenia czułości defektoskopu w odniesieniu do wad, w szeregu rur spawanych, których złącza nie wykazywały większych niejednorodności [50], wykonano wady testowe. Były to frezowane kanałki wzdłużne o różnych głębokościach oraz otwory wywiercone przez jedną ściankę rury. Prezentowane niżej wyniki badań dotyczą rur z wadami testowymi w postaci otworów. Średnice otworów wynosiły 0,5 mm; 0,8 mm; 1,0 mm; 2,0 mm i 3,0 mm. W rurach wykonano także grupy otworów. W celu otrzymania poszczególnych wad testowych o równomiernej średnicy, otwory wiercono poprzez metalową płytkę.

Jako czułość defektoskopu należy przyjąć zdolność wykrywania wad materiałowych, rozumianą jako odstęp sygnałów, wywołanych przez wady od poziomu tła.

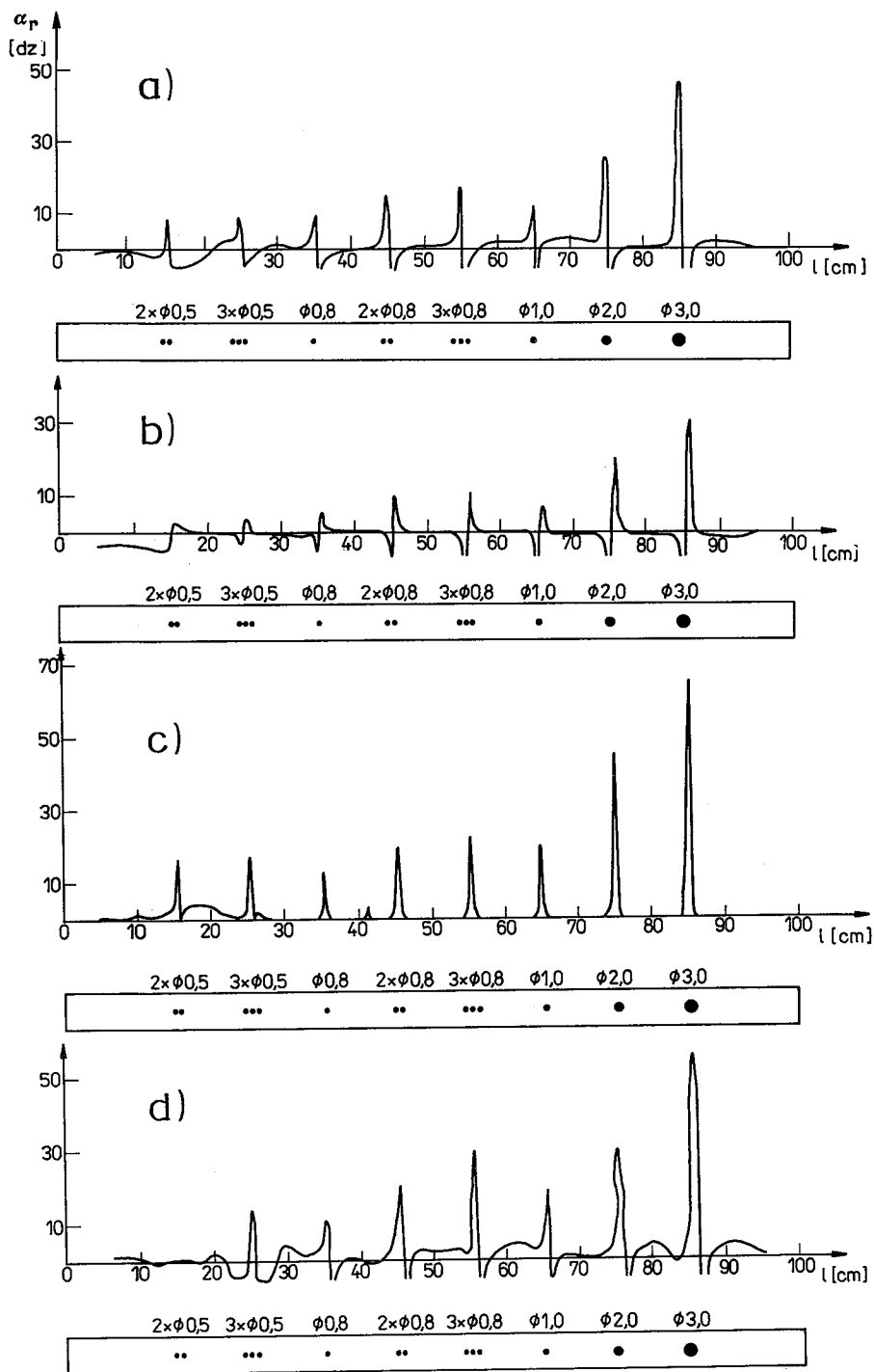
Na rys. 6 przedstawiono wskazanie miernika wyjściowego defektoskopu w odniesieniu do sekcji różnicowej przetwornika w funkcji długości dla rur miedzianej i mosiężnych, z wadami testowymi. Średnice nominalne rur, zgodnie z produkowanym w Polsce asortymentem, wynosiły 6,5 mm; 10 mm; 16 mm i 18 mm. Pomiary wykonano przy stałym wzmocnieniu. Na podstawie przytoczonych wyników badań można stwierdzić, że wykrywane są 2 otwory średnicy 0,5 mm (rura mosiężna średnicy 6,5 mm — rys. 6a i rura mosiężna średnicy 16 mm — rys. 6c) oraz 3 otwory średnicy 0,5 mm (rura miedziana średnicy 10 mm — rys. 6b) i rura mosiężna średnicy 18 mm (rys. 6d). We wszystkich rurach wykrywano pojedynczy otwór o średnicy 0,8 mm. Tak więc czułość defektoskopu odpowiada wymaganiom normy [50].

## 8. WADY NATURALNE ZŁĄCZ SPAWANYCH RUR MIEDZIANYCH I MOSIĘŻNYCH

Studia nad obiektami badanymi — rurami miedzianymi i mosiężnymi — spawanymi wzdłużnie wraz z nieniszczącymi badaniami defektoskopowymi i niszczącymi badaniami metalograficznymi pozwoliły na określenie rodzajów wad, występujących w spoinach tych rur. Stwierdzono, że w złączach rur występuje wiele różnych wad, a mianowicie: pęcherze gazowe, przetop na części grubości ścianki, pęknięcia, przepalenia oraz defekty powierzchni, kształtu i rozmiarów złącz [59]. Należy podkreślić duże nasilenie występowania wad. Obszerny materiał odnoszący się do przeglądu wad, występujących w złączach rur stalowych, zawarto w [2]. Przedyskutowano tam również przyczyny występowania wad złącz rur.

Na fot. 7 — fot. 10 przedstawiono fotogramy 14 zglądów metalograficznych poprzecznych wybranych przekrojów złącz rur miedzianych i mosiężnych. Przekroje rur otrzymano poprzez cięcie wolnoobrotową piłą diamentową. Po zainkludowaniu otrzymanych w ten sposób próbek, wykonano zglądy poprzez szlifowanie przy użyciu papierów korundowych o granulacji 3  $\mu\text{m}$  i 1  $\mu\text{m}$  oraz tlenku glinu o granulacji od 100 do 800 i polerowanie — przy użyciu past diamentowych o granulacji 0,25  $\mu\text{m}$ . Próbkę trawiono 5-cio procentowym roztworem kwasu siarkowego [1]. Fotogramy zglądów wykonywane były przy użyciu mikroskopu NEOPHOT-2 firmy Carl Zeiss Jena. Fotografowano próbki przed trawieniem w celu wyodrębnienia wad makrostruktury (pęknięcia, pęcherze gazowe, itd) oraz po trawieniu — dla zbadania mikrostruktury złącz, tzw. strefy wpływu ciepła oraz materiału macierzystego [58]. Na wszystkich przedstawionych fotogramach zewnętrzna ścianka złącz rur znajduje się u góry.

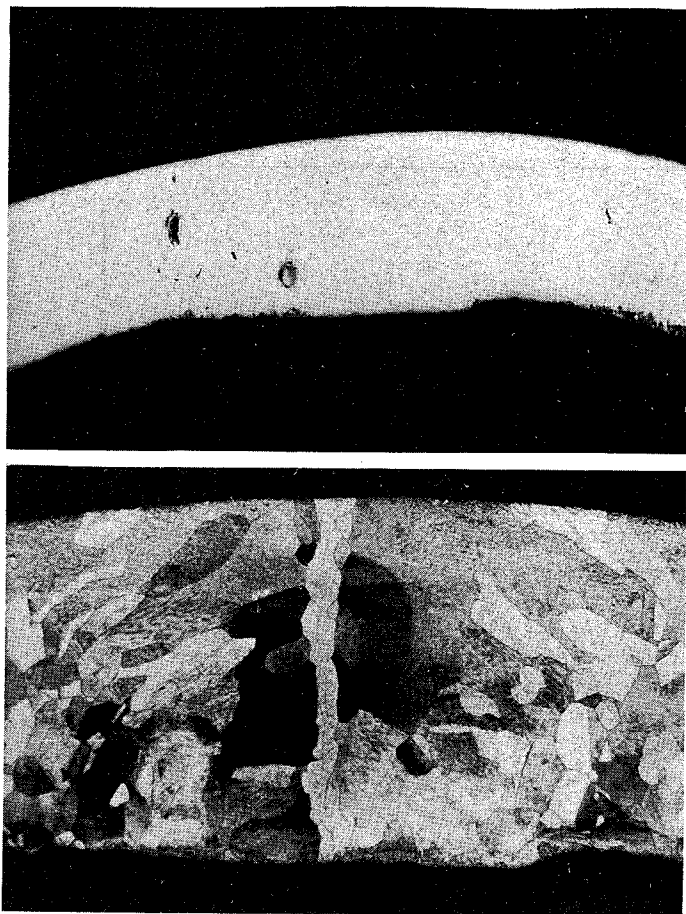
Na fot. 7 przedstawiono fotogramy złącz rur miedzianych. Złącze rury pokazane na fot. 7a ma kształt zbliżony do prawidłowego (nadlew grani nie przekracza 10% grubości ścianki rury). Widoczne są sporej wielkości pęcherze gazowe. Złącze rury przedstawione na fot. 7b, charakteryzuje się znaczną szerokością, wyciekami w grani spoiny (tj. od strony wewnętrznej ścianki rur) oraz dużą liczbą niewielkich pęcherzy gazowych. Na fotogramie zglądu trawionego widoczne są duże ziarna w spoinie oraz



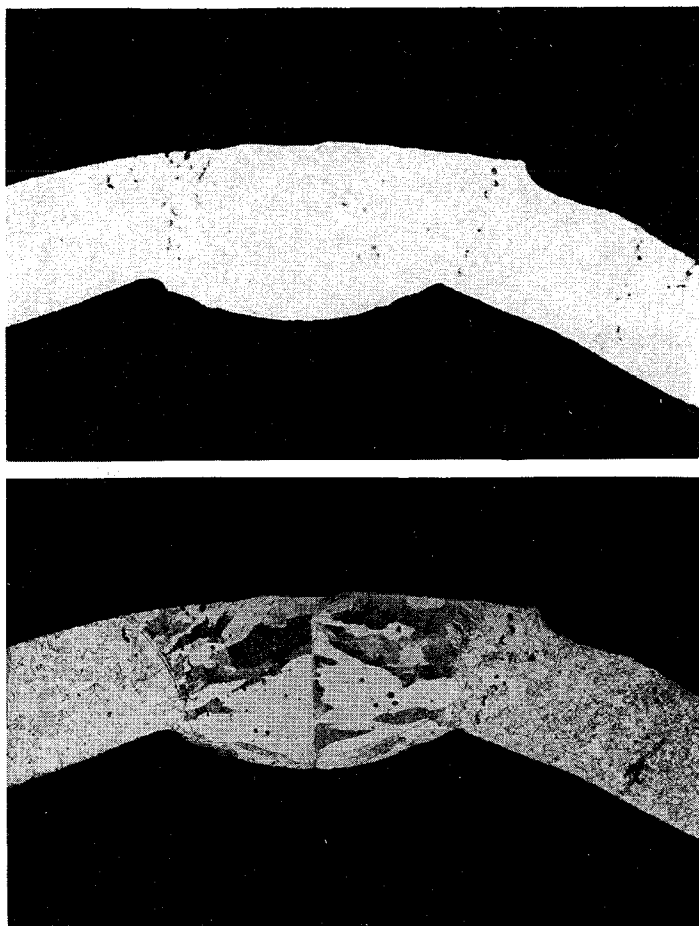
Rys. 6. Wskazania defektoskopu dla rur z wadami testowymi

rozrost w strefie wpływu ciepła — w porównaniu z drobnoziarnistą strukturą materiału macierzystego — niepodgrzewanego podczas procesu spawania. Widoczna jest wyraźnie linia wtopienia (tj. linia dzieląca spoinę od strefy wpływu ciepła). Charakterystyczny jest układ ziaren w spoinie, świadczący o kierunkach odprowadzania ciepła podczas krzepnięcia złącza. Wąski pas ziaren po środku spoiny zakrzepł ostatni (fot. 7a i fot. 7b). W złączu rury miedzianej, pokazanym na fot. 7c, występują pęcherze gazowe. Jedna z krawędzi taśmy jest przesunięta względem drugiej, tj. występuje tzw. uskok. Widoczne są typowe dla miedzi kryształki bliźniacze [78].

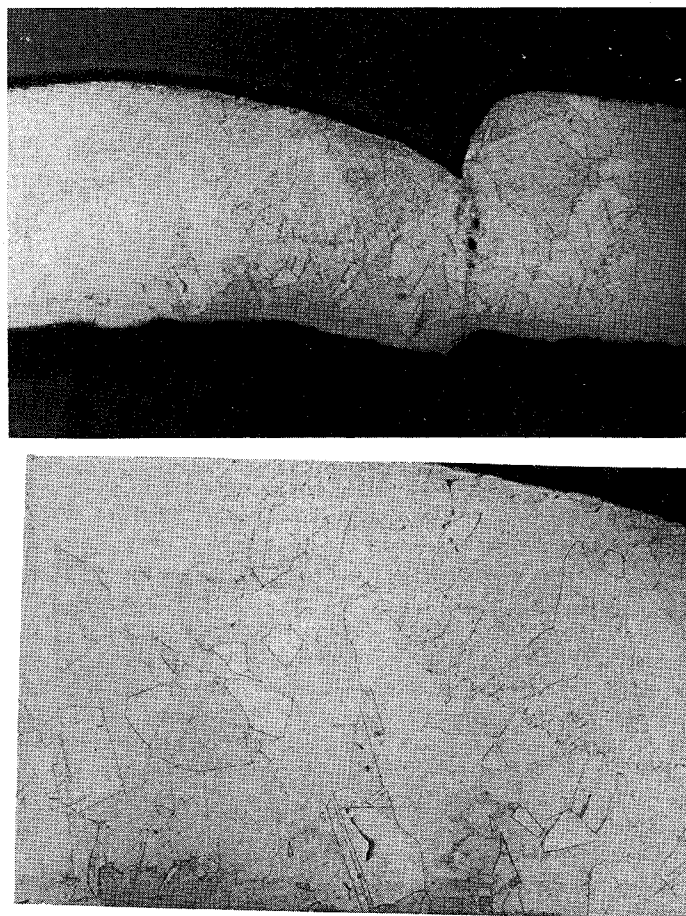
Na fot. 8 pokazano złącza rur miedzianych oraz złącze rury mosiężnej. Pęknięcia złącz rury miedzianej (fot. 8a i fot. 8b) zalegają odpowiednio na połowie i znacznej części grubości ścianki rury; pęknięcia pozostałych złącz, tj. spoiny rury mosiężnej oraz rur miedzianych (fot. 8c — fot. 8e) są przelotowe. W złączu rury mosiężnej (fot. 8c) występuje charakterystyczne dendrytyczne (słupkowe) ziarno, wskazujące na dużą szybkość odprowadzania ciepła w pierwszej fazie krzepnięcia spoiny. Taka struktura



Fot. 7. Fotogramy zglądów metalograficznych złącz rur miedzianych (zglądy nie trawione i trawione); a) rury średnicy 8 mm i grubości ścianki 0,5 mm; powiększenie  $\times 65$  i  $\times 130$



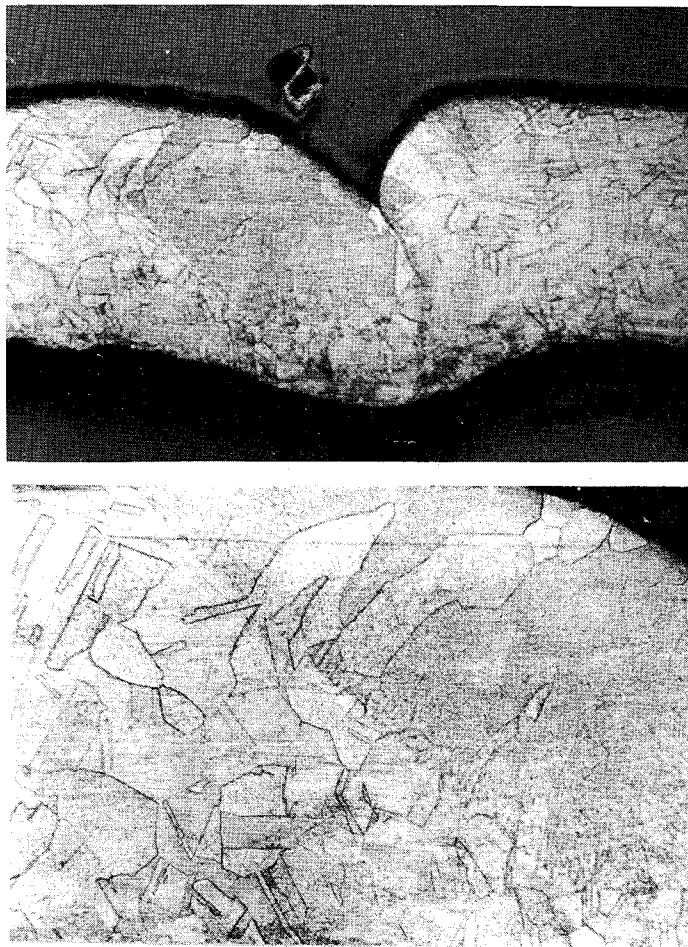
Fot. 7. b) rury średnicy 16 mm i grubości ścianki 1 mm; powiększenie  $\times 25$



Fot. 7. c) rury średnicy 16 mm i grubości ścianki 0,6 mm; powiększenie  $\times 60$  i  $\times 120$

złącza występuje wówczas, gdy rury mosiężne nie zostają poddane wyżarzaniu normalizacyjnemu (rury miedziane, jako wyroby z materiału jednofazowego wyżarzaniu takiemu nie podlegają). Złącza rur miedzianych przedstawione na fot. 8d i fot. 8e, charakteryzują się obecnością pęknięć o znacznej szerokości. Typowa jest gruboziarnista struktura utworzonej kropli (fot. 8e). W obu złączach wystąpiło nadmierne przetopienie materiału.

Na fot. 9 pokazano złącza rur mosiężnych oraz rury miedzianej, odznaczające się wadami kształtu. Złącze rury mosiężnej, przedstawiono na fot. 9a, charakteryzuje się niesymetrycznym wyciekem w grani oraz nieregularnością lica (od strony ścianki zewnętrznej rury). Występuje ponadto znacznej wielkości pęcherz gazowy. W złączu rury mosiężnej, pokazanym na fot. 9b, występuje znaczny wyciek w grani, nieregularność lica oraz grani. Struktura złącza w znacznej części jest dendrytyczna. Ze



Fot. 8. Fotogramy zgładów metalograficznych złącz rur miedzianych oraz rury mosiężnej (zgłady nietrawione i trawione) a) rury miedzianej średnicy 16 mm i grubości ścianki 0,6 mm; powiększenie  $\times 60$  i  $\times 120$



Fig. 8. b) rury miedzianej średnicy 16 mm i grubości ścianki 0,6 mm; powiększenie  $\times 60$

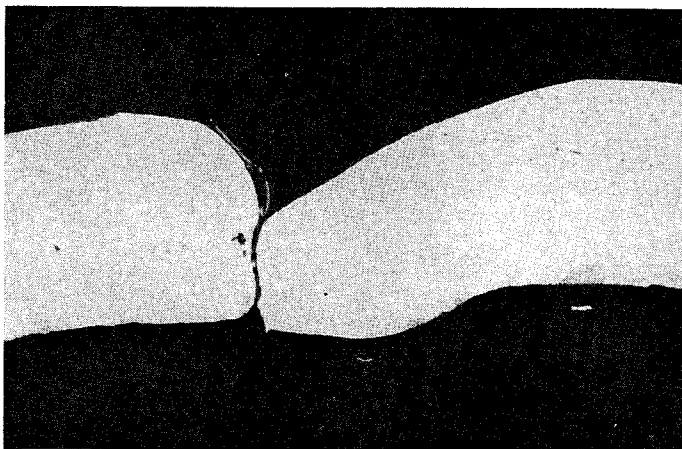


Fig. 8. c) rury mosiężnej średnicy 16 mm i grubości ścianki 0,6 mm; powiększenie  $\times 65$

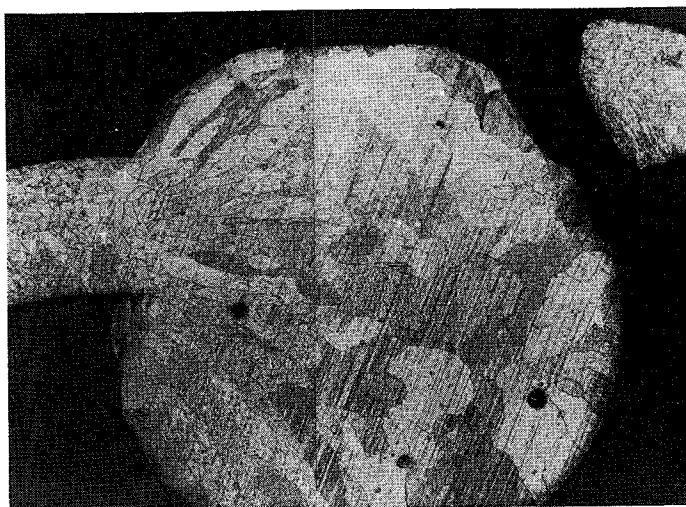
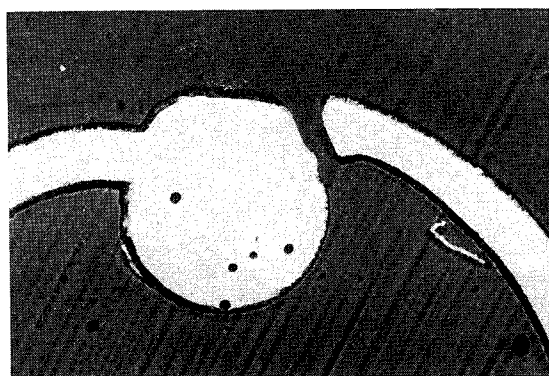
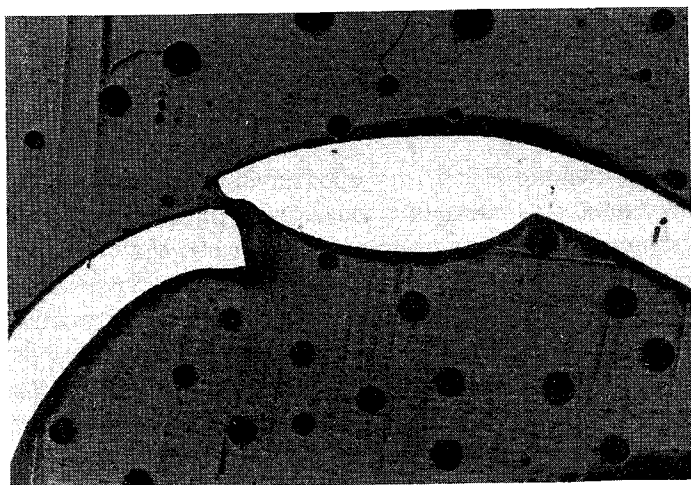
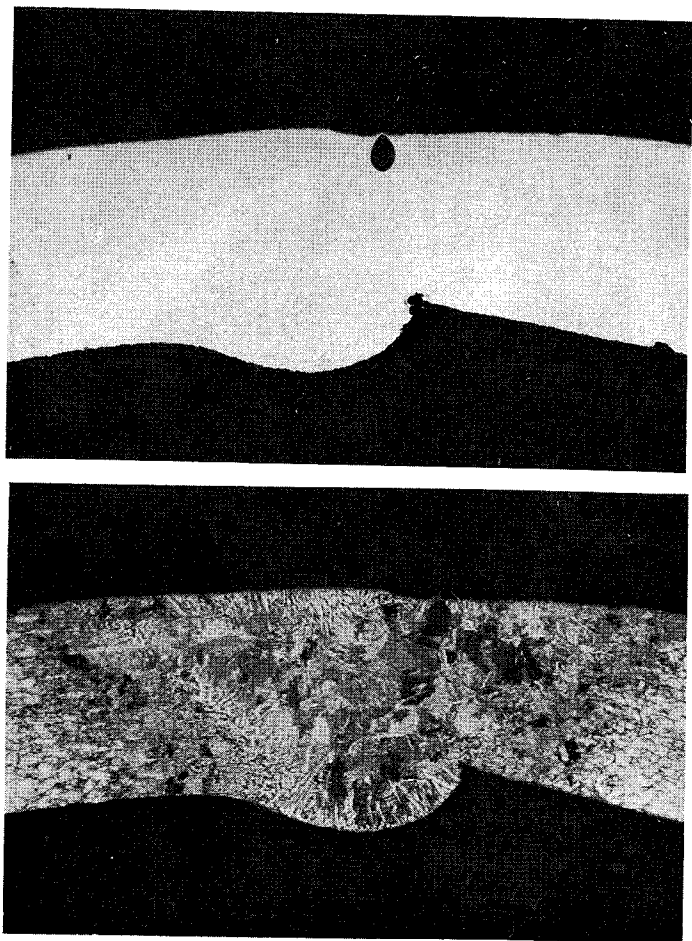


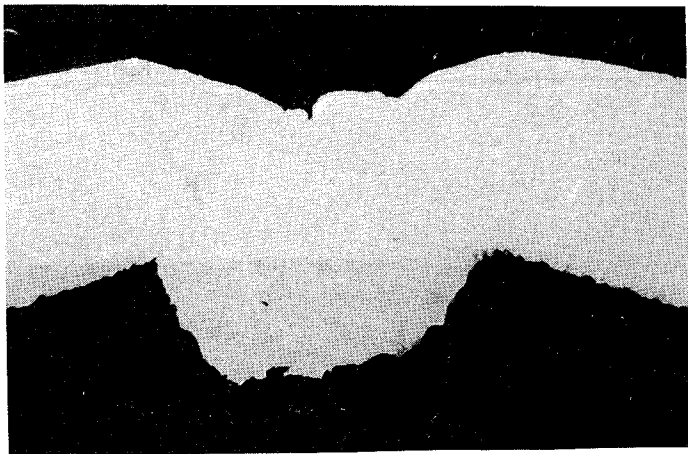
Fig. 8. d, e) rur miedzianych średnicy 8 mm i grubości ścianki 0,5 mm; powiększenie  $\times 25$

względem na przetopienie zbyt dużej ilości materiału, spoina opadła wskutek działania sił grawitacji. W złączu rury miedzianej (fot. 9c) wystąpił znaczny wyciek. Widoczne są pęcherze gazowe. Złącze rury mosiężnej, przedstawione na fot. 9d, ma znaczną szerokość, wskutek czego opadło pod działaniem sił grawitacji. Występuje w nim struktura dendrytyczna [1, 78].

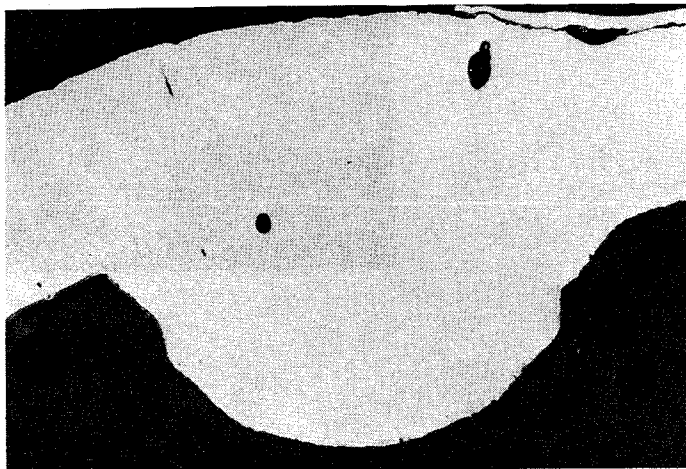
Na fot. 10 przedstawiono złącze rury miedzianej oraz rury mosiężnej. W wyniku przetopienia znacznej ilości materiału w złączu rury miedzianej (fot. 10a) utworzona została kropla, zajmująca znaczną część przekroju rury. Materiał zebrany w kropli, podobnie jak w przypadku złącza, pokazanego na fot. 8e i fot. 9c został zgromadzony w badanym przekroju na pewnym odcinku rury. Charakterystyczna jest struktura materiału kropli (gruboziarnista) oraz struktura materiału od strony przeciwległej do kropli ścianki rury — występuje „płynięcie” materiału. Spoina rury mosiężnej, przedstawiona na rys. 10b, oderwała się i w płaszczyźnie prezentowanego przekroju nie



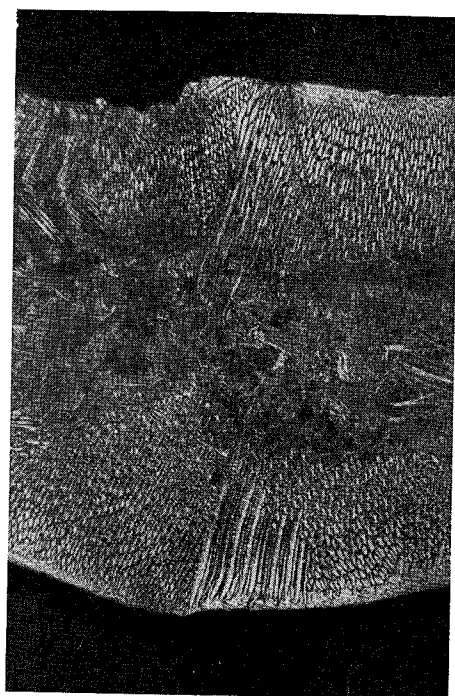
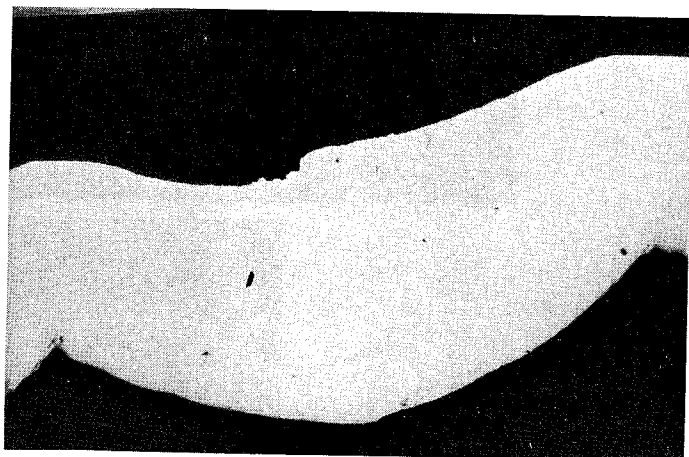
Fot. 9. Fotogramy zglądów metalograficznych złącz rur mosiężnych oraz rury miedzianej (zglądy nietrawione i trawione) a) rury mosiężnej średnicy 16 mm i grubości ścianki 0,6 mm; powiększenie  $\times 65$



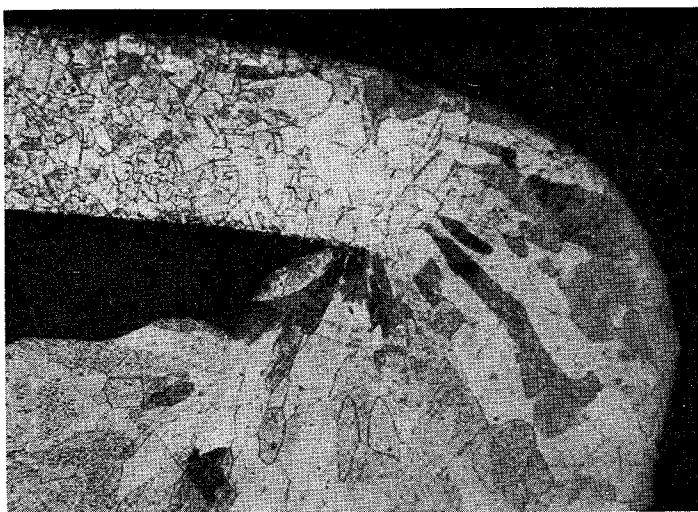
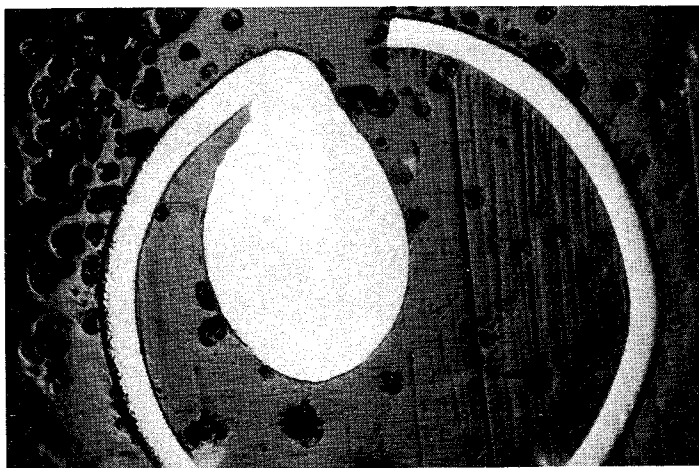
Fot. 9. b) rur mosiężnych średnicy 10 mm i grubości ścianki 0,65 mm; powiększenie  $\times 65$  i  $\times 130$



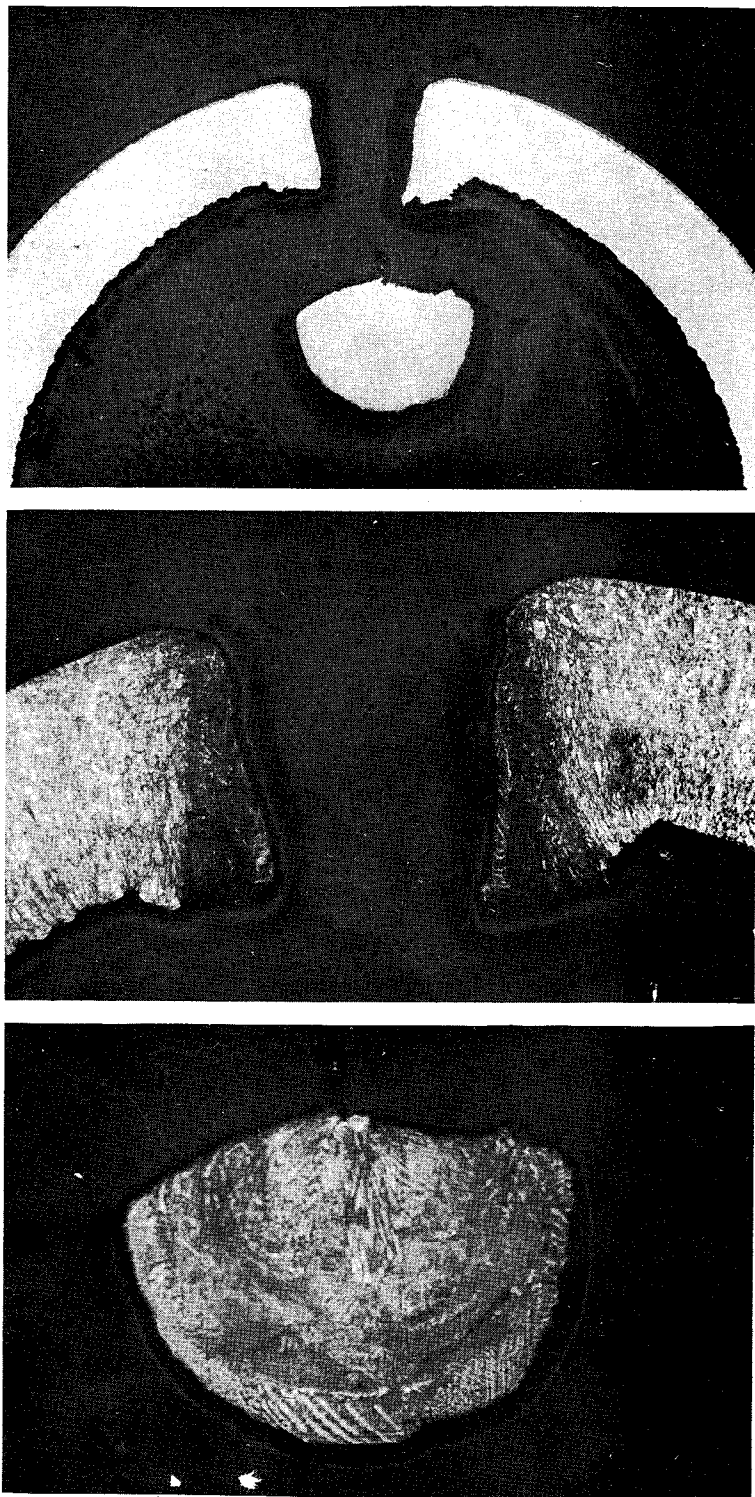
Fot. 9. c) rury miedzianej średnicy 8 mm i grubości ścianki 0,5 mm; powiększenie  $\times 65$



Fot. 9. d) rur miedzianych średnicy 10 mm i grubości ścianki 0,65 mm; powiększenie  $\times 65$  i  $\times 130$



Fot. 10. Fotogramy zglądów metalograficznych złączy rury miedzianej i mosiężnej (zglądy nietrawione i trawione) a) rury miedzianej średnicy 8 mm i grubości ścianki 0,5 mm; powiększenie  $\times 12,5$ ,  $\times 65$ ,  $\times 125$



Fot. 10. b) rury miedzianej średnicy 8 mm i grubości ścianki 0,6 mm; powiększenie  $\times 25$  i  $\times 65$